Don’t Let Your Data Get SMACked: Introducing 3-D Data Management

As social, mobile, analytics and cloud continue to disrupt business, organizations need a new approach to data management that supports real-time data requirements, ever-changing and fast-growing data volumes and the need for new types of transactions.
Executive Summary

Taming data management across the enterprise has always presented challenges and complexities, from the time-consuming and costly process of creating a “gold copy” of master data to making the right data available at the right time. But now, with social, mobile, analytics and cloud technologies (the SMAC Stack™), many traditional data management techniques and technologies simply cannot keep pace with the disruptive demands for agility, speed, new types of fast-growing data and innovative ways of working and transacting.

A new “three-dimensional” approach to data management can resolve the issues and challenges posed by the SMAC Stack. This approach addresses today’s need for real-time delivery of the right data, to the right place, in the right form by:

• Replacing static data integration approaches with a five-layer framework that we call the “GRiD™” which uses triggers to ensure all functions of the enterprise get the right data at the right time.

• Ensuring data fidelity by building it into business processes rather than incorporating it as an infrastructure element.

• Using the emerging approach of crowdsourcing as an affordable way to mine large volumes of fast-changing data to produce relevant and timely insights.

This white paper explores the new demands on data management posed by the SMAC Stack; describes a new data management approach that responds to these challenges; and provides examples of this approach in action.
Fifth Wave of IT

The business world has experienced several waves of corporate IT architecture foundations in the past several decades, moving from mainframes, to minicomputers, to distributed PCs, to the Internet and now the SMAC Stack. We call this the “fifth wave” of IT, and like its predecessors, SMAC is causing significant disruption and displacement, even while it drives new levels of productivity. (For more on SMAC’s disruptive force, see our white papers, “Don’t Get SMACked: How Social, Mobile, Analytics and Cloud Technologies are Reshaping the Enterprise” and “Code Rules: A Playbook for Managing at the Crossroads.”) In response, businesses need to update their data management strategies to ensure they can continue delivering valuable insights based on both their own data stores and new types of unstructured data, as well as enable the business to respond to new ways of working and transacting.

We propose a new approach for organizing, transporting, analyzing and managing data assets, one that operates across three dimensions.

The four components of the SMAC Stack are all impacting data management, each in its own way:

- **Social media:** Across the globe, people are flocking to social media. According to a recent Pew Research survey, almost three-quarters (72%) of online adults in the U.S. use social networking sites. All of this data — from employment history and connections, to relationship status, to personal and professional interests — can be mined for insights across three dimensions: customer behavior, networking potential and brand sentiment. Organizations today must redo their information roadmaps to include these important data sources, and build in functionality to separate the signal from the noise. The goal is to drive business insights from these rich data stores to increase their margins and establish mindshare.

- **Mobile:**Few technologies have impacted the personal and work lives of people as much as mobile devices. As mobile devices proliferate, they are changing the habits of consumers and employees, who are becoming increasingly reliant on connectivity and personalized, location-based service and content. Customers today also expect to engage with companies across multiple channels (mobile, social, Web, phone, in-store), and they expect these interactions to be continuous and consistent across channels. As this behavior takes hold, businesses need to migrate toward an approach that can analyze real-time information on consumer location, behavior and needs to enable delivery of the right product or service, across all channels.

- **Analytics:** With the explosion of content available, it has become critical to enable decision-making through advanced analytics, along with the ability to handle “big data,” or fast-changing, fast-growing, mixed types of data. Analytics is needed to unearth insights into customers, products and locations, using data from multiple sources, including unstructured data (both internal and external), structured data and third-party information. The challenge for companies is enabling the best analytics capability without exorbitant cost.

- **Cloud:** Traditional data management platforms relied on consolidating data within the in-house infrastructure. But the vast volumes of data, varied content types and complex integration scenarios generated by social and mobile platforms re-
quire a more real-time and integrated approach, utilizing a cloud infrastructure. Companies need to discern what to move to the cloud and what to keep in house, as well as build a strategy that helps them move to the new architecture.

3-D Data Management

Traditional precepts of data management — such as approaching data quality as a linear function or consolidating data in a static data store — are an inadequate response to the cumulative changes brought about by big data and the demands for localized and personalized transactions across channels. We propose a new approach for organizing, transporting, analyzing and managing data assets, one that operates across three dimensions (see Figure 1):

- **Integration**: Delivering a trigger-based, interconnected data integration platform to ensure the right data is available at the right time and place.
- **Data fidelity**: Ensuring data quality is delivered and is measurable in terms of impact to business processes.
- **Data analytics**: Utilizing faster, cheaper analytics to deliver the right insights.

**Dimension 1: Integration**

Enterprises are grappling with the basic problem of acquiring and delivering information in a timely manner and in the right format, particularly because of the size, speed and multiple formats of data that exist today.

A traditional data management approach calls for businesses to consolidate data from systems throughout the enterprise into a data store or data mart. In other cases, companies have created point-to-point integrations among systems with the most critical data. However, neither of these approaches delivers results at the speed required by businesses today.
In order to deliver data expediently, economically and in a timely fashion throughout the business, we propose a five-layer data integration architecture we call the GRiD, short for Get Right Data and similar in construct to an electrical grid (see Figure 2).

The GRiD’s five layers include:

- **Operational systems:** These systems are where data gets created and is consumed. This is also where triggers are initiated, signaling data updates to the rest of the GRiD.

- **Triggers:** When data is updated (such as an address change), a trigger notifies the relevant areas of the GRiD.

- **Metadata:** Triggers are translated into something meaningful so other parts of the GRiD can decide the most appropriate action to take.

- **Composite data services:** This is an external-facing layer that allows the outside world to access data within the GRiD, at an abstracted level that makes the most sense to them.

- **Process orchestration:** This is where processes that interact with the data are combined in a way that makes sense to the business.

In a real world example, a healthcare benefits manager used the GRiD to deliver business process efficiencies integrated with master data. In doing so, it enabled a comprehensive information infrastructure that sends out notifications to other systems and processes when data is created or updated. The GRiD reduced billing errors by 12%, reduced manual operating costs by 6% and improved customer service scores by 2%.

For a healthcare benefits manager, the GRiD reduced billing errors by 12%, reduced manual operating costs by 6% and improved customer service scores by 2%.

---

**Understanding the GRiD**
Dimension 2: Data Fidelity

While the GRiD focuses on the transport of data, the second dimension focuses on the quality, or fidelity, of data and the quality of the insights that can be generated by the data. This dimension includes governance (providing the right framework for the data), data volatility (dealing with the dynamism of data over time) and data integrity (ensuring accuracy, consistency, completeness and veracity).

Establishing data quality has historically been considered a challenging and time-consuming endeavor. For one thing, quality is contextual – it changes over time and according to the process using it (i.e., data that is considered valid for one process may not be considered as such when used by another). Additionally, changes to data are not always available at the point of capture, and processes ensuring high-quality and trusted information can be somewhat more of an art than a science.

To overcome these challenges, we propose pursuing data quality efforts within the context of larger business processes rather than as a foundational element. Whereas most data management initiatives work to establish data quality within the infrastructure (i.e., establishing a “gold copy” of the data to be shared with all business processes), we believe data quality should be built into individual business processes themselves. This can be accomplished by establishing “process hubs,” which include four components:

- **Define** the end-to-end business process and subprocesses.
- **Identify** the linkage between the process and the data (e.g., “create, read, update and delete,” or CRUD), as well as the process and data pain areas.
- **Document** process-related performance metrics and use these to establish data metrics.
- **Establish data governance** by using the business rules surrounding the process.

**Whereas most data management initiatives work to establish data quality within the infrastructure, we believe data quality should be built into individual business processes themselves.**

By using process hubs in place of global hubs, business can realize several benefits, such as easier establishment of qualitative success metrics; fewer variances in data definitions; and better boundaries around the scope of data management.

A major financial services company took a step-by-step approach to building an enterprise master data platform, using individual process hubs, supported by a common governance framework across the enterprise. The hubs were targeted at specific processes, including risk profiling, client onboarding, “know your customer” (KYC), Internet commerce and securities management.

This approach resulted in multiple quantifiable benefits. For one, it helped the company improve its risk-scoring-related errors by 30% and reduce corresponding costs. The financial service provider also improved customer satisfaction scores during the onboarding process by 50% and reduced compliance errors during the know your customer/anti-money laundering (KYC/AML) process by 40%. Lastly, it improved the results of its marketing campaign across various channels by 12%.
**Dimension 3: Data Analytics**

Analytics has catapulted to the top of most companies’ priority lists; however, this capability is far easier to get wrong than right. It is one thing, for instance, to generate a data-driven report and quite another to pull actionable insight from that data. Additionally, while analytics has traditionally focused on the present or the past, today’s more competitive business environment requires organizations to prepare for the next business event. This demands predictive analytics, which is not as easily supported by most systems today. Lastly, the ability of most enterprises to provide relevant and timely insights lags far behind the petabytes of data they are storing. With burgeoning data volumes, traditional approaches to analytics will result in exponential budgetary expenditures, particularly in terms of the systems and storage to store the data.

In order to control costs while obtaining actionable insights, we propose an approach already being taken by many forward-thinking companies: crowdsourcing, in which analytics models are built by talented individuals found in the public domain. Companies set up these projects as coding challenges or competitions and issue a monetary reward for the best submission. Industry leaders such as Netflix and Allstate are prime examples of how crowdsourcing can benefit business objectives at a fraction of traditional IT costs. While crowdsourcing is still an emerging technique, it is an approach that has shown much promise in controlling costs and delivering powerful insights.

**Looking Forward**

In the future, additional dimensions may be added to this forward-thinking data management approach. We can foresee the value, for instance, in "social data management," in which organizations would be able to identify influential individuals on social media and interact with these individuals to encourage specific patterns of behavior.

For now, however, it is enough for businesses to realize the shortcomings of current approaches to data management and begin to think in terms of a three-dimensional approach. There is no question that we’re moving to a fifth wave of IT corporate architecture, and it’s time to optimize the technologies of our time.

**Footnotes**


About the Authors

Sriraj Rajaram is Cognizant’s Enterprise Data Management Practice Lead. As a trusted advisor to customers in various industry segments, he specializes in information strategy, data governance, master data and data integration. Over 15 years, he has helped pioneer industry-leading solutions such as MDM-in-a-Box™ and Customer-in-the-Cloud™. He can be reached at Sriraj.Rajaram@cognizant.com | Linkedin: http://www.linkedin.com/pub/sriraj-rajaram/1/243/4b5.

Naveen Sharma is Cognizant’s Master Data Management Practice Lead, which is part of Cognizant’s Enterprise Data Management Practice. He has more than 17 years of software development and data management experience, 13 of which have been focused on MDM. Naveen specializes in helping customers define their MDM roadmap and strategy and contributes to the creation of industry-leading MDM solutions at Cognizant. He can be reached at Naveen.Sharma@cognizant.com | Linkedin: http://www.linkedin.com/in/4naveen.
About Cognizant
Cognizant (NASDAQ: CTSH) is a leading provider of information technology, consulting, and business process outsourcing services, dedicated to helping the world's leading companies build stronger businesses. Headquartered in Teaneck, New Jersey (U.S.), Cognizant combines a passion for client satisfaction, technology innovation, deep industry and business process expertise, and a global, collaborative workforce that embodies the future of work. With over 50 delivery centers worldwide and approximately 164,300 employees as of June 30, 2013, Cognizant is a member of the NASDAQ-100, the S&P 500, the Forbes Global 2000, and the Fortune 500 and is ranked among the top performing and fastest growing companies in the world. Visit us online at www.cognizant.com or follow us on Twitter: Cognizant.

About Cognizant’s EAS
Cognizant’s Enterprise Application Service (EAS) takes a customer-focused approach to helping clients enable end-to-end business processes and deploy intuitive, new-age enterprise platforms to drive efficiency, innovation and profitable growth.

World Headquarters
500 Frank W. Burr Blvd.
Teaneck, NJ 07666 USA
Phone: +1 201 801 0233
Fax: +1 201 801 0243
Toll Free: +1 888 937 3277
inquiry@cognizant.com

European Headquarters
1 Kingdom Street
Paddington Central
London W2 6BD
Phone: +44 (0) 207 297 7600
Fax: +44 (0) 207 121 0102
infouk@cognizant.com

Continental Europe Headquarters
Zuidplein 54
1077 XV Amsterdam
The Netherlands
Phone: +31 20 524 7700
Fax: +31 20 524 7799
infonl@cognizant.com

India Operations Headquarters
#5/535, Old Mahabalipuram Road
Okkiyam Pettai, Thoraipakkam
Chennai, 600 096 India
Phone: +91 (0) 44 4209 6000
Fax: +91 (0) 44 4209 6060
inquiryindia@cognizant.com

© Copyright 2013, Cognizant. All rights reserved. No part of this document may be reproduced, stored in a retrieval system, transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the express written permission from Cognizant. The information contained herein is subject to change without notice. All other trademarks mentioned herein are the property of their respective owners.