



DATA DRIVEN SERVICES PORTFOLIO



IN-MEMORY DATA FABRICS

J DATA WAREHOUSE & MASTER DATA MANAGEMENT

DATA MANAGEMENT CONSULTING



FACT SHEET

IN-MEMORY DATA FABRIC CONSULTING IN-MEMORY DATA FABRICS FOR DATA INTENSIVE APPS AND I.O.T. PROCESSING

SCALE PROCESSING WITH IN-MEMORY DATA FABRICS

IN-MEMORY DATA FABRIC CONSULTING

□ APACHE GEODE (open-source GEMFIRE)

PIVOTAL GEMFIRE

APACHE IGNITE (open-source GridGain)

SCALE WITH IN-MEMORY DATA FABRICS CONSULTING

Today's data-intensive applications can no longer be serviced by disk-based storage systems. Even vertically scalable RAM-based solutions – *due to in-server resource contention and their singular purpose* – cannot deliver the low-latency I/O and advanced compute capabilities required.

Distributed **In-Memory Data Fabrics** change all of that. Horizontally scalable with an ability to add nodes on demand, data fabrics linearly scale to hundreds of nodes while providing strong semantics for data locality and affinity-datarouting to reduce redundant data noise. The result? Unprecedented speed and scale to dataintensive applications. Some data fabrics also offer powerful features such as a compute grid; high performance ACID transactions; stream and complex event processing; support for SQL querying and distributed joins; Spark and Hadoop acceleration; and so on.

Planning, deploying and developing against a data fabric can be somewhat daunting. To begin, your organization will need to select a platform based on the merits; and then weigh the pros and cons of selecting the open-source or enterprise edition of that platform.

From P.O.C. to architecture to development, we fully engage along the adoption cycle to help!

IN-MEMORY DATA FABRIC CONSULTING

Whether performing real-time analytics across a I.O.T. device collection; or responding to events from a PUB/SUB bus, an In-Memory Data Fabric will provide the low-latency capabilities that you need. We offer design, develop and deploy consulting services for the *GemFire*, *Geode* and *Ignite* platforms.

BUSINESS USE-CASE UNDERSTANDING

To best leverage the functionalities of a particular inmemory data fabric product, it's important to understand the business use cases that it will support. These drive the selection of a particular topology and caching mode; enabling maximum enterprise benefit from the solution. Unlike master-slave platforms, inmemory data fabrics make use of peer-to-peer architectures, so use case understanding is particularly important.

□ ARCHITECTURAL CONFIGURATION SELECTION

Based on enterprise usage patterns identified (there can be multiple), available machinery, use case SLAs, and mission criticality (such as business continuity), architectural decisions can be made, including:

- Cluster topology
- Caching strategy
- Cluster sub-groups (logical clusters in a fabric)
- Data model
- Multi-datacenter replication
- Asynchronous support
- Resilience

□ INSTALLATION AND CONFIGURATION

Each product comes with particular roadmaps and requirements for installation; but apart from those, deployment details will include some or all of the following considerations: Native versus containerized installations; Roles-based access and security integration; Event and alert NOC integration; Server access, including as root; and basic testing of the data fabric once fully deployed.

□ ENTERPRISE INGEGRATION

Whether to provide 2nd level caching to avoid expensive disk access and database calls; or to enable high-performance transaction processing for eCommerce shopping carts; or to facilitate real-time event notification and processing for credit card fraud identification and risk calculation; or to provide an in-memory compute grid to aid I.O.T. and Spark analytics; programmatic enterprise integration will be necessary to harness the power of these platforms. During this phase, we offer best-practice developer consultants to do this.

OPERATIONAL ADJUSTMENTS AND TUNING

Distributed computing works great until it doesn't. You'll be scratching your head in bewilderment when it doesn't: Is it a JVM pause issue? A thread pool issue? Operational analytics instituted during the installation phase can help, but sometimes you'll need deeper troubleshooting. We can help!



IN-MEM DATA FABRIC CONSULTING for better business outcomes! web: www.prismalytics.io email: connect@prismalytics.io





D/W & MDM CONSULTING SERVICES

ENABLE B.I. USERS WITH SUPERIOR DATA WAREHOUSE AND MDM, PLATFORMS

BEST PRACTICE, D/W & MDM ENGAGEMENTS

D/W & MDM CONSULTING SERVICES:

□ DATA WARESOUSE (D/W)

❑ MASTER DATA MANAGEMENT (MDM)

WHAT CAN YOU ACHIEVE WITH D/W & MDM CONSULTING?

FACT SHEET

Data Warehouse and Master Data Management platforms are *key long-term business investments* for exploiting and solving datacentric challenges. Yet there is no single, predetermined, correct architecture for either. Many variables factor into the relative success or failure of a solution, and that is where D/W & MDM CONSULTING SERVICES can help.

Other than requiring a parallel query engine that can read large amounts of data, a **DATA WAREHOUSE** conforms to no particular standard, and does not implement a particular application. In fact, it's design goals are business organic: capture business-specific objectives and processes, including data source dependencies, into a data model that enables meaningful analytics. D/W structures are difficult to populate and maintain, and take a long time to construct; so methodical planning and execution is a must.

MDM *reference-data* systems are also driven by business-specific sources, but bring their own challenges. Here, attributes of business *dimensions* are assigned complex rules to determine which of multiple business sources should be consulted for their value. This leads to sprawling networks of sources and destinations, connected by various platinum-priced ETL machinery; usually too brittle to even maintain.

DW & MDM CONSULTING SERVICES

Data Warehouse and **MDM** platforms are complex undertakings. Not surprisingly, companies with a long-term view of these key assets have the most successful outcomes. Our engagements encapsulate the best-practice deployment and iterative-maintenance steps that lead to similar success:

□ CAPTURE AND ANALYZE BUSINESS DRIVERS

These include: Cross-functional business discovery; Business requirements; Identifying core business processes; Data sources; Gap analysis; and so on.

CONCEPTUAL DATA MODEL DESIGN

This includes: Determining the subjects that will be expressed as fact tables and the dimensions that will relate to the facts; Identifying KPIs for each business process and deciding the format to store the facts in; Deciding on consistent units of measure; Relating the dimensions to the KPIs; and so on.

□ DATA-TRACKING DURATION (GRANULARITY)

Data warehouses have a goal of providing consistent analytical performance against an ever-growing set of historical data, so an archiving strategy is crucial. Data structures within a warehouse are set to retain data at various levels of granularity. The granularity must be consistent within a structure, however each structure can have it's own granularity; and then all of them can be related through shared dimensions.

As data ages, it can be summarize and stored in other structures with less granularity. For example, one could store the first 2-years of data with a daily-grain; store

3-to-5 year old data with a weekly-grain; and store data older than that with a monthly-grain. As data ages, they would be migrated to successively coarser-grained tiers. *By planning these stages, one can design analysis tools to work with the changing grains based on data age.*

□ WAREHOUSE DATA -VS-MASTER DATA

WAREHOUSE DATA and MASTER DATA serve different purposes, have different characteristics and are serviced by different systems. Among other differences, the former is historical in nature and used for analytical purposes; while the latter is always current and used to ensure accurate upto-date entity data. In some architectures, data warehouses can consume MDM data as one of it's sources. Separating these two data types and use cases is important.

TECHNICAL DRIVERS

These include determining: Infrastructure requirements (on/off premise); Data integration/logistics/ETL design and development requirements; Data Quality requirements; Deployment and U.A.T. planning; and so on.

MDM BUY -VS- BUILD

Platinum-priced MDM solutions can have poor ROI and be too brittle to maintain. Build-to-fit alternatives, for example using Python and a RDBMS, can often achieve better technical and business characteristics. Talk to us about this.



D/W & MDM CONSULTING for better business outcomes! web: www.prismalytics.io email: connect@prismalytics.io

DW & MDM CONSULTING SERVICES

Full-cycle, long term engagements will typically cover the following first-iteration phases:

- ✓ CORPORATE-LEVEL AND L.O.B. ENGAGEMENT
- ✓ BUSINESSES & REQUIREMENTS DISCOVERY
- ✓ BUSINESSES PROCESS DISCOVERY
- ✓ GAP ANALYSIS
- ✓ PROJECT PLAN DEVELOPMENT

- ✓ PERSONNEL PLAN (INITIAL AND ONGOING)
- ✓ INFRASTRUCTURE PLAN (OFF/ON PREMISE)
- ✓ DESIGN AND DEVELOPMENT PLAN
- ✓ DEPLOYMENT AND USER ACCEPTANCE TEST PLAN
- ✓ CUSTOMER EDUCATION PLAN

The following diagram depicts the main interactions possible with clients over the course of the first iteration of a successful project; as well as subsequent ongoing iterations to capture and incorporate new business developments. Although it illustrates a **D/W** engagement, **MDM** progressions are similar.









DATA MANAGEMENT CONSULTING MAXIMIZE DATA-ROI BY IMPROVING DATA MARGENENT AND UTILIZATION

BEST PRACTICE, DATA MANAGEMENT CONSULTING

DATA MANAGEMENT CONSULTING:

- **BC / DR DATA REPLICATION**
- **DATA MIGRATION**

□ INFORMATION LIFECYCLE MANAGEMENT (ILM)

DATA MANAGEMENT CONSULTING FOR BETTER DATA-ROI

FACT SHEET

Data is the lifeblood of every business and its owners have three critical responsibilities to it: **managing**, **protecting** and **leveraging** it.

Managing and protecting data are generally corporate-level responsibilities, and include: protecting against loss corruption; it or implementing roles-based access to it; migrating wrapping information lifecycle it: and management (ILM) practices around it. Leveraging data is line-of-business BI centric. and often requires aggregating and duplicating data, along with associated data-logistics in order to satisfy analytic needs. These activities create even more data and data-locations to manage.

Complicating matters is the fact that **raw** and **aggregate** data should be classified separately for **traceability** and **recoverability** purposes: aggregate data should be versioned and tagged to reflect the raw data assets from which they are derived; and, in turn, those raw assets should be tagged as immutable so that derived data can always be reconstituted from them.

Every organization benefits from independent audits of their data standards and consumption. Whether it's to determine if your BC/DR dataset is correct or if your ILM practices can be improved; our experienced consultants can help identify gaps and implement remediation steps.

DATA MANAGEMENT CONSULTING

By it's nature, data management is a complex task that is never complete. Organizations that are best at it employ **continuous review and improvement cycles**. We offer a suite of services that maximize data-ROI, by improving data management, protection and utilization. Engagement services include:

BC/DR DATA REPLICATION

After identifying the mission-critical subset of data required for the continuance-of-business at remote locations, we can implement a mixture of synchronous and asynchronous storage-array, database and O/S-level data replication, based on recovery time and recovery point business objectives (RTO/RPO). Planning includes I/O quiescing of "consistency groups" to ensure data and application-level consistency at the destination. **ISO 22301 & 22313** are consulted in this engagement for best practices.

DATA MIGRATION

A variety of business-technical conditions can trigger the need to mass-migrate data from one storage platform to another: *platform generation upgrade; switching to a new vendor platform; relocation to a different price/performance tier in line with the value data has to the business; relocation from on-premise to off-premise storage platforms; and so on.* Samevendor migrations can be accommodated using platform-native capabilities; while cross-vendor migrations necessitate the use of a gateway operating *system.* Copy sequence, consistency and throughput are important, and our experts take you through the entire migration planning and orchestration lifecycle.

□ INFORMATION LIFECYCLE MANAGEMENT (ILM)

ILM is a comprehensive approach to managing the flow of an information system's data and associated metadata from creation and initial storage, to the time when it becomes obsolete and is deleted or archived. ILM involves all aspects of dealing with data, starting with user practices (rather than simply automating storage procedures, as with HSM). Beyond data age and frequency of access, ILM enables complex criteria for storage management – including assignable policies that can decline in weight over time or remain constant – and determine the storage capability-tier that it should be assigned to. Combined with virtual access-path management applications, shifting of data assets between storage tiers can be made transparent to end users.



DATA MANAGEMENT CONSULTING for better business outcomes!web:www.prismalytics.ioemail:connect@prismalytics.io