Improving Media & Entertainment Performance with Big Data
Architect’s Guide and Reference Architecture Introduction

ORACLE ENTERPRISE ARCHITECTURE WHITE PAPER | FEBRUARY 2015
Disclaimer

The following is intended to outline our general product direction. It is intended for information purposes only, and may not be incorporated into any contract. It is not a commitment to deliver any material, code, or functionality, and should not be relied upon in making purchasing decisions. The development, release, and timing of any features or functionality described for Oracle’s products remains at the sole discretion of Oracle.
Table of Contents

Executive Summary ........................................ 1
Key Business Challenges .................................. 4
Where to Find Business Cases that Justify Projects .... 6
Establishing an Architectural Pattern .................. 9
IT Operational ETL Efficiency ............................. 12
Oracle Products in the Information Architecture .... 13
Additional Data Management System Considerations .. 15
Extending the Architecture to the Internet of Things ... 17
Keys to Success .............................................. 20
Final Considerations ......................................... 22
Executive Summary

The ability to access, analyze, and manage vast volumes of data while rapidly evolving the Information Architecture is increasingly critical to Media and Entertainment (M&E) companies looking to improve business efficiency and performance. Content distribution and consumption related challenges continue to dominate the industry. Operational efficiency and understanding customer’s preferences and experiences remain keys to success. Anticipating demand is important for more efficient content management, revenue generation and overall profitability. Media and Entertainment companies are shifting to content centric work to embrace a customer centric world. Consumers are becoming increasingly selective in what they want to view and purchase. Consumers are moving away from subscription services and gravitating towards purchasing content and then using it on any device at any time. This behavior is clearly shaping how the content is delivered and related business models. Top drivers to business are projected to be (1) multi-platform distribution, (2) increasing the customer base and profitability, and (3) growing advertising revenue.

In order to deliver content for a variety of platforms/devices, it is critical that organizations establish platforms that can scale up and effectively enable this multi-distribution model. Organizations need to enable their data warehouse, analytics, and content delivery architectures to a point where they can deliver differentiated and personalized services across a range of devices.

M&E companies are also facing competitors with a variety of go-to-market strategies. Increasingly, they are using the internet as a means to deliver content. Where they also own physical entertainment facilities (such as theme parks and cruise lines), they are seeking to gain a single view of the customer across all of these touch points.

Understanding the customer enables M&E companies to better predict the success of new forms of content and delivery. The data sets available to do this analysis are larger and more diverse and management of this data is more complex. Analysis of this data can lead to better understanding of what content and offerings drive the highest profitability.

A well documented example (in the New York Times and elsewhere) that many M&E companies are trying to mimic is Netflix’s creation of “House of Cards”. Netflix knew from the analysis of their customers their likes and dislikes, including ratings that indicated that many were fans of political dramas. Netflix also knew that they liked the actor Kevin Spacy and that they liked David Fincher’s work. All of this together helped frame the highly successful creation of “House of Cards”. The impact to Netflix is that targeted programming is bringing them more subscribers and greater revenue, profitability, and market share.
Media and Entertainment companies have long gathered customer data to understand who watched their content including the demographics and socio-economic traits. Ratings measured which content hit the target and which content missed expectations. Now, Big Data and analytics can enable a better understanding of the consumer as content is created to better ensure success.

Utilizing a variety of data sources can help M&E companies gain a better understanding of future customer demands. These data sources can include:

» Social Media
» Web browsing patterns
» Traditional enterprise data from operational systems
» Data from data aggregators (Nielsen, etc.)
» Advertising response data
» Demographic data
» Historical M&E data

As noted earlier, some M&E companies also have a physical presence in theaters, theme parks, cruise lines, and retail. This drives the need to analyze many additional data sources such as customer traffic, staffing, supply chain, and logistical information to provide an optimal customer experience and increase value and revenue from all of the properties.

The rate at which this data is generated is rapidly increasing leading to higher rates of consumption by the business analysts who crave such information. This increase in data velocity and sources naturally drives an increase in aggregate data volumes. Business analysts want more data to be ingested at higher rates, stored longer and want to analyze it faster. “Big Data” solutions help to enable M&E companies to meet these requirements.

This paper provides an overview for the adoption of Big Data and analytic capabilities as part of a “next-generation” architecture that can meet the needs in the dynamic Media and Entertainment market.

This white paper also presents a reference architecture introduction. The approach and guidance offered is the byproduct of hundreds of customer projects and highlights the decisions that customers faced in the course of their architecture planning and implementations. Oracle’s advising architects work across many industries and government agencies and have developed standardized methodology based on enterprise architecture best practices. Oracle’s enterprise architecture
approach and framework are articulated in the Oracle Architecture Development Process (OADP) and the Oracle Enterprise Architecture Framework (OEAF).
Key Business Challenges

Media and Entertainment companies historically used data warehouses and business intelligence tools to report on operations, customer subscriptions, and analyze scheduled programming or other entertainment popularity. Today, customers can choose when and what they want to watch or read using DVRs (Digital Video Recorders) and advanced set top boxes, and/or downloading content on-demand over the internet. M&E companies are struggling with legacy business models, legacy infrastructure, and information silos as they seek to gain an understanding as to how well their content and offerings are matching customer demand. For many, this understanding is critical to being able to convince advertisers that the advertising rates they are being charged are fair.

Further complicating these efforts, the rapid development of social networks is changing how content is curated and recommended. Every user of social media is now a curator to friends and the larger community beyond. Distribution platforms are deepening their integration with social networks to respond to this demand for sharing. For example, Facebook’s integration with Spotify gives any user the ability to share his or her musical tastes with friends. The Sony PlayStation 4 demonstrates that social network integration must be core to the gaming experience.

By deploying Big Data Management Systems that include traditional data warehouses and newer data reservoirs (featuring Hadoop and/or NoSQL Databases), much broader types of data can be analyzed to ensure that the business can become more agile. Key business challenges include improving customer intimacy and operational efficiency, and providing information as a paid service.

Improving Customer Intimacy

Media and Entertainment companies seek to differentiate themselves by providing content and entertainment that their consumers want to consume. To do this they must have a better understanding of their customer, including what, how, and when customers select content and offerings. Also, understanding customer’s extended network of family and friends preferences and relating them to the individual preferences of the customer will help develop a unique customer profile resulting in improved overall personalization of content creation and delivery.

For example, it is no surprise that individuals sharing playlists with friends, "liking" a film on Facebook or reviewing a book on Amazon have all become mainstream ways to influence consumer choices—and are now at par with or even above traditional sources of recommendations. Some believe that more than half of video gamers in Western countries choose games based on their friends’ social network recommendations. Understanding these influencers and targeting them with special perks can lead to increased popularity and revenue for content and offerings.

Recommendations are also being automated through aggregation and analysis of behavior data — the “Big Data”. A growing number of consumers rely on algorithmic recommendation engines to help them determine what to read or watch or attend.

Improving Operational Efficiency

Operators of entertainment venues such as theme parks and cruise lines are experimenting with electronic bracelets and sensors to track customers’ movements. Benefits to the customers include faster access to events and attractions and more personalized experiences. Benefits to the entertainment companies include a better understanding of what their customers want and more revenue per customer.
Also desirable is to understand staffing and facilities utilization trends, and other costs of running the business, including maintenance and supplies. Predictive analytics are sometimes used to predict whether current trends will continue and possible implications.

Content providers are re-examining the right delivery mechanism for their services. Linked to their understanding of customers, certain age groups might be more likely to view content via the internet with possible operational benefits tied to decreased carrier costs and lower production costs. Both traditional and emerging delivery channels must co-exist due to changing demographics and increasingly tech-savvy customers. Multi-device content delivery is expected to dominate the market place across the various customer demographics.

Providing Information as a (Paid) Service

Whereas M&E companies traditionally relied on others to be data aggregators, many are now looking for opportunities to sell data and / or invent new lines of business based on their ability to deliver content of value. Subscriptions are often offered for customer data for reuse in targeted marketing. News, weather and sports information is increasingly for sale in a variety of forms to other businesses (such as insurance companies, retailers, and other entertainment companies).
Where to Find Business Cases that Justify Projects

Many existing business capabilities can be enhanced when more and varied data becomes part of the Information Architecture. The following figure illustrates many of these common where a single customer view is established in M&E companies.

Figure 1: The Single Customer View in Media & Entertainment companies

IT organizations at M&E companies typically work with their lines of business to build solutions that deliver the following when defining Big Data projects:

1) **Single View of Customer across Content Delivery Platforms & Offerings**: As M&E companies broaden their content delivery platforms and offerings, they must understand how customers interact with each to fully understand customer value and opportunities to up-sell and cross-sell. They must understand how customers change their preferences for certain platforms over time. Such an understanding also enables setting of justifiable rates for advertising across platforms and pricing for individual and combinations of platforms. Marketing offers and campaigns can be more accurately targeted. The value of the customer to the company can grow, even if they lose interest in certain delivery platforms, content, or other offerings over time.

2) **Customer Retention through Personalized Content**: Understanding what customers want to consume and how they want to consume the information enables the building of customized offerings leading to better customer retention and additional sources of revenue. M&E companies can deliver the content in the format that is most desirable to the customer by understanding how they are gaining access to content and offerings, their sentiment about the experience as expressed on social media, the amount of time they are viewing / listening to content or engaged at entertainment venues, and impactful changes in their lifestyle (such as the addition of children) on the
kind of content that is of interest. They also use this data to determine the business opportunity linked to the
acquisition of new content.

3) Advertising Revenue Growth through Better Targeting: By understanding customers better, how they consume
content and services, and what advertisements attract follow-on activities, M&E companies can better establish the
value of their brands and offerings leading to increased revenue from advertisers. They can also offer services and
subscriptions to a potentially valuable customer database for the advertisers to mine themselves.

4) Timely Retirement of Assets of Limited Value: M&E companies need to understand when the right time is to
retain assets by archiving content or retiring increasingly unprofitable delivery channels and offerings. This requires
detailed knowledge of customer trends and the linkage between various offerings that at first might not be apparent.

5) IT operational efficiency: Not unique to M&E companies and rarely driven from the lines of business (but a
possible reason for embarking on extended architectures that include Hadoop) is the need to move data staging and
transformation to a schema-less platform for more efficient processing and leveraging of IT resources. IT
operational efficiency is often difficult to prove but is sometimes an initial justification that IT organizations gravitate
toward when deploying these types of solutions.

On the next page, we show a table that summarizes several typical business challenges in M&E companies and the
opportunity for new or enhanced business capability when adding new analytic capabilities.
<table>
<thead>
<tr>
<th>FUNCTIONAL AREA</th>
<th>BUSINESS CHALLENGE</th>
<th>OPPORTUNITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Media / Entertainment Marketing</td>
<td>Obtain the most benefit per dollar of marketing spend</td>
<td>Improve market share</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Improve brand image &amp; loyalty</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cross-sell / up-sell across channels</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Improve ad / promotion responses</td>
</tr>
<tr>
<td>Advertising Revenue</td>
<td>Demonstrate the value of the media channel / entertainment to existing &amp; potential advertisers</td>
<td>Increase ad volume and revenue</td>
</tr>
<tr>
<td>Content Creation, Acquisition, and Programming</td>
<td>Create or purchase and schedule the most compelling content desired by customers for an efficient multi-platform distribution</td>
<td>Increase ad rates</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Retain advertisers by moving to most appropriate channels</td>
</tr>
<tr>
<td>Content Archiving and Retirement</td>
<td>Archive, retire, or sell content and offerings when it no longer is profitable</td>
<td>Maintain existing viewer / customer loyalty and share</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Replace old content with more compelling content</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Maintain / grow ad revenue</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Manage infrastructure cost (e.g. storage)</td>
</tr>
<tr>
<td>Entertainment Venue Management (theme parks, cruise lines, etc.)</td>
<td>Optimize customer experience while growing margins</td>
<td>Route customers more efficiently</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Optimize seating / video / audio experience</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Improve / replace underperforming offerings</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cross-sell / up-sell offerings</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Increase attendance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Achieve positive sentiment</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Optimize staffing, facilities, refreshments, etc.</td>
</tr>
</tbody>
</table>
Establishing an Architectural Pattern

The following figure illustrates key components in a typical Information Architecture. Data is acquired and organized as appropriate and then analyzed to make meaningful business decisions. A variety of underlying platforms provide critical roles. Management, security and governance are critical throughout and are always top of mind in M&E companies. These components are further described in the “Information Architecture and Big Data” whitepaper posted at http://www.oracle.com/goto/ea.

How do we determine which of these components should be part of the architecture to meet the needs of a specific organization or company? If we create an information architecture diagram, and trace the data flow from the sources to the application (end-user), we can build a logical configuration of the components to support the functions.

The first step in defining a future state architecture is documenting the current state, its capabilities and any functional gaps. Typically a current state data warehouse environment might look something like Figure 3.
The first gap that typically has to be closed is a need to provide a more agile reporting and analysis environment where new data and ad-hoc reports are needed on an ongoing basis. Information and data discovery engines can provide this type of capability. When information discovery is incorporated into the architecture it would look something like the illustration in Figure 4.

![Figure 4: Typical Introduction of Information Discovery](image)

Now that we’re better able to analyze the data we have, the next step would be to explore bringing in new data and new data tapes. These data sets might be internal, 3rd party, structured, unstructured or of unknown structure. When storing data of unknown structure, the most efficient way to store data sets is often in a Hadoop-based data reservoir. Initially, such projects are often considered experimental in organizations and therefore they might be independent efforts separated from the traditional environments, as illustrated in Figure 5.

![Figure 5: Typical Early Hadoop Environment separate from the Data Warehouse](image)
The profile of the data such as how it is acquired, how it should be formatted, the frequency of updates and quality of the data will help us put the right technology in place best suited for the particular situation. We need to understand whether real-time or batch processing is appropriate. We should understand the periodicity of processing required based on data availability. Below is a partial list of the characteristics that should be considered:

- **Processing Method** – prediction, analytics, query, ad-hoc reports
- **Format and Frequency** – external data feeds, real-time, continuous or periodic on-demand
- **Data Type** – web/social media, machine generated, human generated, biometric, legacy or internal, transactional
- **Consumer Application** – Web Browser, Intermediate processes, Enterprise Application

When business value is found in analyzing data in a Hadoop-based data reservoir, lines of business generally begin to see a need to link data there to historical data stored in their data warehouse. For example, a business analyst might want to compare historical transactions for a shipment stored in the data warehouse to sensor data tracking that shipment in the data reservoir. Various linkages are often established as pictured in Figure 6.

![Figure 6: Integration of Hadoop Infrastructure and Data Warehouse](image)

We also added something new to Figure 6, a real-time analytics and recommendation engine. In many situations, the latency inherent in the data movement pictured above means that the recommendation from analysis would come too late to take action in near real-time. A way around this is to perform periodic advanced analytics in the data reservoir and/or data warehouse and provide updates to a real-time recommendation engine that becomes more fine-tuned through self-learning over time.
IT Operational ETL Efficiency

In Figure 6, you might have noticed a line pointing from the transactional sources to the Hadoop cluster. This is to illustrate a popular ETL alternative, leveraging Hadoop as a data transformation engine.

Let's now consider the type of data typically stored in today's data warehouse. Such warehouses are typically based on traditional relational databases using a “schema on write” data model. The data sources can vary, but the structure of the data is determined before the data is imported into the data warehouse. In the example below there are two data sources. These two data sources go through an ETL process to prepare the data to be loaded into the warehouse.

Extending the architecture can enable a more agile workflow by incorporating data sets for which there is not rigid structure. This data model is best defined as “schema on read”. That is, we store the data without the traditional ETL processing, as we don’t know exactly how we want to access the data. In the example below we are using multiple data sources with varying structures.

These two environments should not be separate and unique. Building an integrated Information Architecture that can handle data sets of known structure as well as unknown structure enables us to augment the capabilities of existing warehouses as well as leverage data center best practices that are already in place.
Oracle Products in the Information Architecture

In Figure 9, we illustrate how key Oracle products could fit in the generic architecture diagram previously shown.

While Oracle can provide a more complete integrated solution, many organizations mix and match products from a variety of vendors. Therefore, such architecture diagrams often show such a mixture of products from Oracle and other vendors.

Defining an Information Architecture is all about linking it to a specific use case. For example, a use case that includes various operational sources used for determining customer experience might look like Figure 10:
The various software capabilities required in a typical architecture might include these Oracle components:

» Oracle Relational Database Management System (RDBMS): Oracle Database 12c Enterprise Edition is designed for performance and availability, security and compliance, data warehousing and analytics, and manageability. Key data warehousing options often include In-Memory, OLAP, the Advanced Analytics Option, and Partitioning.

» Oracle Business Intelligence Enterprise Edition (OBIEE): A business intelligence platform that delivers a full range of capabilities - including interactive dashboards, ad hoc queries, notifications and alerts, enterprise and financial reporting, scorecard and strategy management, business process invocation, search and collaboration, mobile, integrated systems management and more.


» Hadoop Distributed File System (HDFS): A scalable, distributed, Java based file system that is the data storage layer of Hadoop. Ideal for storing large volumes of unstructured data.

» Flume: A framework for populating Hadoop with data via agents on web servers, application servers, and mobile devices.

» Oracle Data Loader for Hadoop: A connectivity toolset for moving data between the Oracle RDBMS and the Hadoop environment.

» ODI: Oracle Data Integrator is a comprehensive data integration platform that covers all data integration requirements: from high-volume, high-performance batch loads, to event-driven, trickle-feed integration processes, to SOA-enabled data services.

» Oracle Enterprise Metadata Management: Data governance and metadata management tool providing lineage and impact analysis, and model versioning for business and technical metadata from databases, Hadoop, business intelligence tools, and ETL tools.

» Endeca: An information discovery tool and engine.

» Oracle Big Data Discovery: A Hadoop-based information discovery tool.

» Oracle Big Data SQL: An optimal solution for querying an Oracle Database on Exadata and combining the results with data that also answers the query and resides on Oracle’s Big Data Appliance.

» ORE: Oracle R Enterprise enables analysts and statisticians to run existing R applications and use the R client directly against data stored in Oracle Database (Oracle Advanced Analytics Option) and Hadoop environments.

» Oracle Enterprise Manager: An integrated enterprise platform management single tool used to manage both the Oracle structured and unstructured data environments and Oracle BI tools.

» Oracle Essbase: An OLAP (Online Analytical Processing) Server that provides an environment for deploying pre-packaged applications or developing custom analytic and enterprise performance management applications.

The software products listed above can be deployed in an integrated environment leveraging these engineered systems:

» Big Data Appliance (BDA): Eliminates the time needed to install and configure the complex infrastructure associated with build-out of a Hadoop environment by integrating the optimal server, storage and networking infrastructure in a rack.

» Exadata: Streamlines implementation and management while improving performance and time to value for Oracle relational database workloads by integrating the optimal server, storage and networking infrastructure.

» Exalytics: Provides an in-memory server platform for Oracle Business Intelligence Foundation Suite, Endeca Information Discovery, and Oracle Essbase.

Obviously, many variations are possible. As illustrated in Figure 10, part of the architecture might include data analyzed in cloud-based solutions such as Oracle BlueKai.
Additional Data Management System Considerations

In defining the Information Architecture, it is important to align the data processing problem with the most appropriate technology.

When considering the choices you have in database management systems to include in an Information Architecture, you might consider if the form of the incoming data or ACID properties or fast data availability is most important. Other considerations should include manageability, interoperability, scalability, and availability. Of course, you should also consider the skills present in your organization.

Some of the various data management technologies in a typical architecture include:

Relational Databases

Typically already in use at most companies, RDBMS’ are ideal for managing structured data in predefined schema. Historically they excel when production queries are predictable. Support of dimensional models makes them ideal for many business intelligence and analytics workloads. They frequently house cleansed data of known quality processed through ETL workloads. Relational databases also excel at transactional (OLTP) workloads where read / write latency, fast response time, and support of ACID properties are important to the business.

These databases can usually scale vertically via large SMP servers. These databases can also scale horizontally with clustering software.

Example RDBMS Product: Oracle Relational Database

MOLAP Databases

Typically used for highly structured data, MOLAP databases are ideal when you know what queries will be asked (e.g. facts and dimensions are predefined and non-changing) and performance is critical. These databases excel at certain business intelligence and analytics workloads.

Example MOLAP Product: Oracle Essbase, Oracle Database OLAP Option

NoSQL Databases

NoSQL databases are without schema and are designed for very fast writes. Often, they are used to support high ingestion workloads. Horizontal scale is most often provided via sharding. Java and Java scripting (JSON) are commonly used for access in many of the commercial varieties.

NoSQL databases are sometimes described as coming in different varieties:

Key Value Pairs: These databases hold keys and a value or set of values. They are often used for very lightweight transactions (where ACID properties may not be required), and where the number of values tied to a key change over time.

Column-based: These databases are collections of one or more key value pairs, sometimes described as two dimensional arrays, and are used to represent records. Queries return entire records.

Document-based: Similar to column-based NoSQL databases, these databases also support deep nesting and enable complex structures to be built such that documents can be stored within documents.
Graph-based: Instead of structures like the previous types, these databases use tree-like structures with nodes and edges connecting via relations.

Example NoSQL Database Product: Oracle NoSQL Database

Distributed File System

Not a database per se as the name would indicate, highly distributed file systems have the advantage of extreme scalability as nodes are added and frequently serve as a data landing zones or data reservoirs for all sorts of data. Read performance is typically limited by the individual node of the “system” when accessing data confined to that node, however scalability to a huge number of nodes is possible driving massive parallelism. Write performance scales well as data objects can be striped across nodes.

The most popular distributed file system used today is Hadoop. Given its role as a data reservoir, it is increasingly a location for performing predictive analytics. SQL access is available via a variety of interfaces though various levels of standards support are offered.

Example Distributed File System Product: Cloudera Hadoop Distribution (featuring the Cloudera Hadoop Distributed File System and other features)

Big Table Inspired Databases

There is an emerging class column-oriented data stores inspired by Google’s BigTable paper. These feature tunable parameters around consistency, availability and partitioning that can be adjusted to prefer either consistency or availability (given these are rather operationally intensive).

A typical use case might be where consistency and write performance are needed with huge horizontal scaling. HBase (deployed on a Hadoop Distributed File System) in particular has been deployed to 1,000 node configurations in production.

Example Big Table inspired Product: Cloudera Hadoop Distribution (Cloudera HBase)
Extending the Architecture to the Internet of Things

Thus far, we’ve focused on the analytics and reporting and related data management pieces of the Information Architecture. Where the Media & Entertainment company manages physical facilities as part of their brand (e.g. theaters, theme parks, cruise lines, retail outlets, etc.), sensors are increasingly providing key data. The “Internet of Things” architecture for data capture, security, and linkage to the rest of the Information Architecture can require additional consideration. The following illustrates a typical footprint:

Figure 11: Connected Devices in an Internet of Things Footprint

Items to the far right of Figure 11 have largely been previously discussed in this paper. Many of the other items pictured are what Oracle typically describes as Fusion Middleware components. For example, much of the sensor programming today takes place using Java. Security is extremely important since most would not want unidentified third parties intercepting the data provided by the sensors. Applications closer to the sensors themselves are often written using Event Processing engines to take immediate action based on pre-defined rules. There are also various message routing, provisioning, and management aspects of such a solution.

Figure 12 illustrates a typical capability map of this architecture for Media & Entertainment:
Sensors are increasingly used to monitor energy consumption in facilities, supplies, and equipment state. Data monitored from smart bands on attendees at venues and their mobile devices are a growing source of information regarding the amount of time spent in lines, spending patterns, and the popularity of attractions and entertainment options. The capture of such data will continue to grow enabling better management of these facilities.

Figure 13 illustrates some of the Oracle products aligned to the previously shown capability map:
Figure 13: Oracle Products aligned to Capability Map
Keys to Success

One of the most significant keys to success in a large project undertaking is to gain alignment between the business needs and goals and with the IT architecture design and deployment plans. Key business sponsors must be engaged and active in all phases.

Methodologies based on phased approaches are almost always the most successful. To start, you’ll need to understand the current state and its gaps so that you can better understand how to build towards the future state. You will need to modify the architecture as business needs change. Therefore, a common method to help assure success is to deploy quickly in well scoped increments in order to claim success along the way and adjust the plan as needed. A complete Information Architecture is never built overnight, but is developed over years from continued refinement.

Figure 14 illustrates such an approach, beginning with defining an initial vision, then understanding critical success factors and key measures tied to use cases, defining business information maps based on output required, linking the requirements to a Technical Information Architecture, defining a Roadmap (including phases, costs, and potential benefits), and then implementing. Of course, an implementation leads to a new vision and requirements and the process continues to repeat. Pictured in the Figure are some of the artifacts Oracle often helps deliver during Enterprise Architecture engagements and Information Architecture Workshops.

Figure 14: Typical Methodology for Information Architecture Projects

Usability needs will drive many of your decisions. Business analysts will likely have a variety of business requirements and possess a variety of analysis and technical skills. They could require solutions ranging from simple reporting to ad-hoc query capability to predictive analytics. You’ll need to match the right tools and capabilities to the right users. One size does not usually fit all. While new features in the data management platforms can provide more flexibility as to where you host the data for such solutions, the data types, volumes and
usage will usually determine the most optimal technology to deploy. A common best practice is to eliminate as much movement of data as possible to reduce latency.

Data security and governance are also a key consideration. M&E patrons want their data to remain private unless they specifically agree to share it. So securing access to the data, regardless of data management platforms, tools, and data transmission methods used, is critical. Data governance needs regarding the meaning of data as well as its accuracy and quality will often require close coordination with and among multiple lines of business.

Finally, as fast time to implementation important to the success of any business driven initiative, you will want to leverage reference architectures, data models and appliance-like configurations where possible. These can speed up the design and deployment and reduce the risk of incomplete solutions and severe integration challenges. Leveraging engineered systems and appliances where possible can simplify the architecture, reduce time to value and improve architecture reliability.
Final Considerations

This paper is intended to provide an introduction to applying Information Architecture techniques for Media and Entertainment companies. These techniques guide the extension of current architecture patterns to meet new and varied data sources that are becoming part of the information landscape. Oracle has very specific views regarding this type of information architecture and can provide even more of the individual components than were described in this paper.

The following diagram provides a conceptual future state that can encompass all types of data from various facets of the enterprise:

![Typical Conceptual Future State Diagram](image)

Figure 15: Typical Conceptual Future State Diagram


The following is a figure from one of the just referenced documents to give an idea as to the level of detail that might be considered around information delivery and provisioning.
Often, the architecture discussion also leads to consideration on where to host and analyze the data (e.g. in the cloud versus on-premise). Aside from security considerations, most M&E companies come to the conclusion that another motivating factor to storing the data on-premise is the volume of data being produced and a desire to minimize network data traffic. In other words, most organizations are coming to the conclusion that it makes sense to analyze the data where it lands. And once it lands, reporting and predictive analytics often take place in the data management system holding the data.

An additional consideration not addressed in this paper is the availability of skills needed by the business analysts and the IT organization. A future state architecture evaluation should include an understanding as to the degree of difficulty that a future state might create and the ability of the organization to overcome it.

The continued change in how consumers consume information, movies, music, television and entertainment as well as the competitive nature of M&E companies assure that those that take advantage of these new data sources to augment what they know about their business will continue to be leaders. They will continue to invent new and better business processes and efficiencies and they will do so by evolving their Information Architecture in an impactful manner. Some will likely leverage their advanced footprints to offer data subscriber networks, thereby going into competition with data aggregators and further monetizing their IT investments.
Oracle Corporation, World Headquarters
500 Oracle Parkway
Redwood Shores, CA 94065, USA

Worldwide Inquiries
Phone: +1.650.506.7000
Fax: +1.650.506.7200

February 2015
Oracle Enterprise Architecture White Paper – Improving Media & Entertainment Performance with Big Data
Author: Art Licht, Robert Stackowiak, Ned Chini, Venu Mantha