A PROJECT MANAGER'S NOTES:

HOW TO PLAN A SUCCESSFUL DATA WAREHOUSE MIGRATION

TIPS AND STEPS TO CONSIDER

BY TERESA HOLLADAY, PMP, CSM

Data warehouse migration is one of the hottest trending topics in Enterprise Data Management.

In today's increasingly data-hungry world, the average household is expected to create enough data in the next three years to fill 318 iPhones annually. As a result, businesses have quickly discovered that regardless of time or cost, their analysts simply won't use data that's unreliable, irrelevant, difficult to access, or slow to reach their users. Data leads to decisions, and the generation of increasingly large amounts of data is creating a business demand for data warehousing and business intelligence.

For a data warehouse migration to be successful, the data needs to be trustworthy, delivered quickly, and be tightly aligned with end-user needs.

The purpose of this article is to give project managers and technical architects a fast, easy, and practical method to plan for a successful project. Although this article covers a data warehouse migration, this method could be applied in other implementations, such as creating a brand new data warehouse.

PART I: WHAT IS A DATA WAREHOUSE MIGRATION?

Data warehouse migration is the transfer of data from old systems to a new repository. It includes:

- Moving an existing data warehouse from one platform to another
- Modernizing or upgrading an existing data warehouse with new and improved data, structure, hardware, or software
- Creating a new data warehouse from a variety of disparate source systems

Various factors drive the desire for migration:

- Business growth, both organic and through mergers and acquisitions
- "The need for speed"
- The prospect of cost savings
- The potential to use data analytics to grow the business
- Standardization across disparate sources or legacy data warehouses

Regardless of the type or reason for the migration, the fact remains that these are large, complex efforts. Knowing how to plan is essential.

PART II: TIPS TO CONSIDER

1

TIP 1: GO AGILE

One benefit of Agile is much-improved user interaction, which is a critical success factor in data migration. Users will see value more quickly and frequently with the iterative approach than with a traditional waterfall path, and will be able to add input on modifications to increase relevance. The increased teamwork, collaboration, and communication in a truly Agile project can dramatically enhance productivity. You can also opt for a blended approach rather than full Agile, depending on which solution makes sense for your project.

We find that even when the corporate culture or nature of the project supports a Waterfall approach, there are many advantages to incorporating Agile values and aspects of an Agile framework.

2

TIP 2: ENGAGE DATA MIGRATION CONSULTANTS FOR KEY ROLES ON THE PROJECT

- In-house expertise on new platforms and tools is likely to be limited
- For this type of custom effort, the in-house team may not have the strategic planning, architecture, and build experience
- “The old way of doing things” may have already created challenges between IT and business users that a new player can help overcome

3

TIP 3: GO OLD SCHOOL, USE A PROJECT MANAGER

(Source: Dave Crolene, “The Agile Data Warehouse”)

The level of technical, logical, and political complexity involved in a data warehouse project goes far beyond other project types. A Project Manager provides the full-time focus required to manage the intricate release planning process and ensures data modeling embraces the complete enterprise – not just the most prominent subject area. With a Project Manager leading the migration, the project also benefits from risk management and facilitation among a broad group of participants.

4

Tip 4: Begin with the End in Mind

Every project should start with a principle from Stephen Covey's book, The Seven Habits of Highly Effective People, "Begin with the end in mind."

The initiation phase of your project plan is more than a task checklist; it's how you start to truly collaborate with both the technical team and the business. The objective is to get a clear picture of the business user's core needs.

Sit down with principal stakeholders to determine their critical success factors, but don't stop there. Meet with technical teams, including database and system administrators, security, and operations, to find their pain points with past implementations. Learn their definition of success.

Your highest level success criteria may be defined as a Data Warehouse that is:

Accurate – provides trustworthy, reliable data corresponding to the source system(s)

Fast – executes both processing and retrieval times within defined windows

Supportable – captures and easily restarts process failures with exceptions

Verifiable – tests established data quality, including reconciliation

Usable – provides reporting ability on today's data, with additional capabilities afforded by historic data

Scalable – expands to meet both an increase in volume and new, unexpected data

Flexible – adapts to rapidly changing business needs

Secure – appropriately masks data to meet data security requirements

Collaborative – combines in-house and project team personnel and facilitates knowledge transfer and data warehouse support ability

Provides Operational Success – processes run consistently and successfully across development, test, and production platforms, and can be executed by in-house users

In this step, be sure to identify new capabilities the system will deliver, such as data quality, data governance, data masking, new scheduling and BI reporting features, new data sources, and historical capabilities that can facilitate trending.

PART III: PLAN TOGETHER USING A DELIVERABLES-BASED WORK BREAKDOWN STRUCTURE (WBS)

Completing a thoughtful, deliverables-based WBS at the beginning of the project is another Critical Success Factor. It helps manage scope and timeline. Stepping through the process helps team members think through tasks and deliverables and surface potential issues. Because technical teams are usually not familiar with this technique, this is an opportunity from the earliest planning stages for the technical architects to work together with the Project Manager to map out a plan.

According to the Project Management Institute (PMI), there are several benefits to using a WBS:

- Decomposition and better control of the project scope
- Improved estimating
- Better control of the project execution
- More accurate project completion verification

The WBS informs rather than becomes your schedule. As you plan with your team, it allows you to visually see the work, or scope, you need to complete the projects. With the right discussion, it helps ensure that you adequately plan for essential activities.

How does it work?

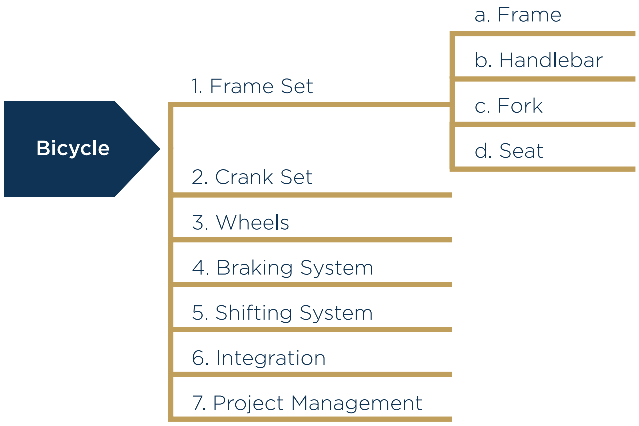
The Project Manager conducts a short series of planning sessions with the team. Beginning with known deliverables and scope, work backwards to break down, or "decompose", the deliverables. This can be done in a Word document, as shown below, in a project planning tool, or in a highly-visual mind-map tool. The goal is to break down each task into what is needed to complete the deliverable.

This classic example is from the PMI Practice Standard for Work Breakdown Structures.
Project Goal: Build a bicycle
Scope, expressed in terms of specific deliverable components:

Bicycle

1. Frame Set (This deliverable is decomposed for the sake of illustration)
    a. Frame
    b. Handlebar
    c. Fork
    d. Seat
2. Crank Set
3. Wheels
4. Braking System
5. Shifting System
6. Integration (bring it all together to create the bicycle)
7. Project Management

The following mind-map WBS shows the deliverable list visually.



As you can see, each deliverable is then broken down, into its component parts. For example, the Frame Set is delineated into Frame, Handlebar, Fork, and Seat.

This process continues until each component is decomposed into small, reasonable work products.

Why choose a deliverables-based WBS?

Because this is a deliverables-based WBS, it identifies specific components to be delivered. Many times, project managers plan based on the activities of a Phase – Analysis, Design, Build, Test, Deploy. But without a deliverables-based WBS, you are only tracking activities, and it would be difficult to know if the actual work products are proceeding on schedule. You also risk missing areas that should be planned within your scope.

Data warehouse migrations are very large projects. In order to deliver on time, it is essential to track against deliverables. For this reason, we recommend the deliverables-based WBS when planning this type of project.

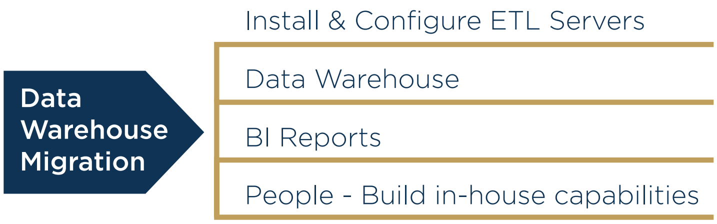
Data warehouse migration example: Let's move from the bicycle example to a data warehouse migration project.

Project Goal: Migrate from a 15-year-old legacy data warehouse to a new data warehouse
Reason: The legacy ETL software is going out of support so new ETL software has been chosen with the database platform remaining the same.
Scope, expressed in terms of highest level project deliverable components:

- Install & configure ETL servers
- Data Warehouse
- BI Reports
- People - Build in-house capabilities for DW/BI to support code and platform

Step 1: Identify the highest-level project deliverables.

The first level of your WBS contains the highest-level project deliverables. If a deliverable is not on this list or would not be decomposed via this list, then it's either not in scope or it should be added to the list. For example: "Install & Configure BI Report Servers" is not on the list. If this deliverable is part of your project, or if it's part of a Statement of Work, it should be part of the WBS.


Data Warehouse Migration
Install & Configure ETL Servers
Data Warehouse
BI Reports
People - Build in-house capabilities

Step 2: Begin to Work Through the Deliverables, Decomposing Each One Down to the Task Level

We'll work through each of the four listed above to illustrate the process. These are high level plans that can be decomposed much further.

Install & configure ETL servers

1. Determine ETL architecture (single or multiple server installation, high-availability, node, grid)
2. Determine hardware specification (size and power)
3. Procure hardware
4. Set up development server
    a. Install and configure OS and security
    b. Install and configure new database to support ETL software
    c. Install and configure new ETL software
    d. Configure backups
5. Create Installation Guide specific to this environment
6. Set up test server
7. Set up production server

Data Warehouse

1. Determine architecture
2. Determine database specification (size and power)
3. Define data architecture standards
4. Complete data model
5. Complete ETL
6. Migrate historical data from old data warehouse

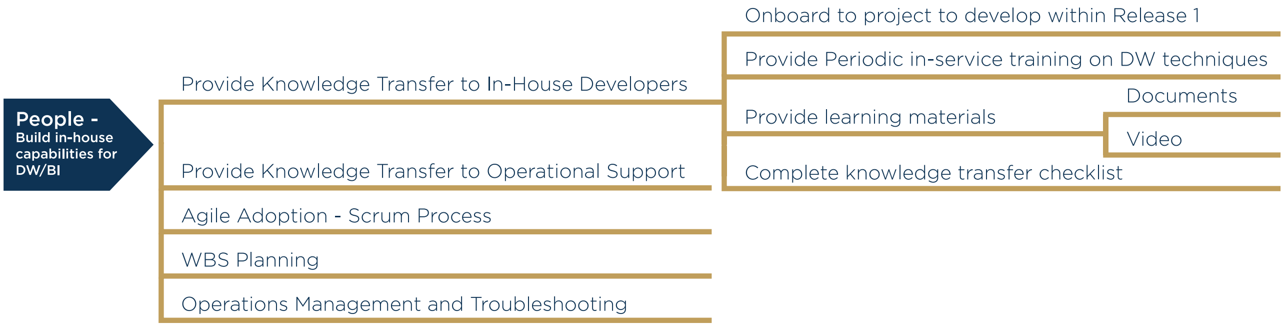
BI Reports

1. Conduct working sessions with users
2. Compile an inventory of highest value reports
3. Determine best approach for BI Report re-tooling
4. Create project plan for BI Report re-tooling
5. Complete BI Report conversion to new DW structure

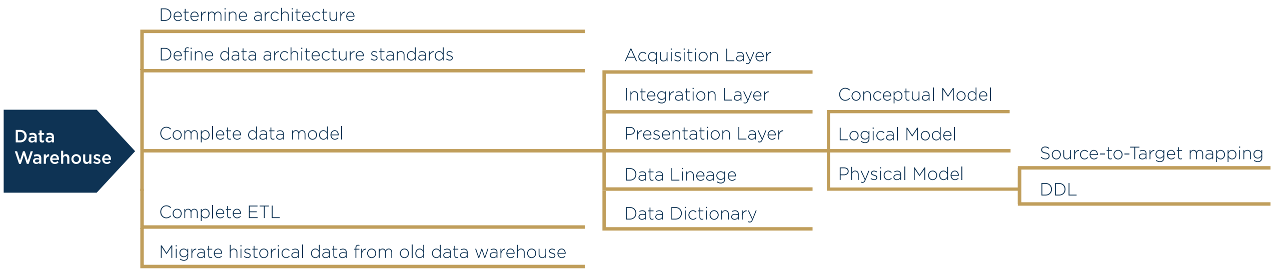
People - Build in-house capabilities for DW/BI to support code and platform

1. Procure software training for in-house Developers
2. Establish process to onboard in-house Developers for hands-on coding
3. Provide knowledge transfer to in-house Developers
4. Provide knowledge transfer to operational support
5. Utilize Scrum process
6. Conduct WBS planning
7. Document operations management and troubleshooting

In a mind-map WBS tool, you can work with your team to easily decompose the deliverables.


People - Build in-house capabilities for DW/BI
Provide Knowledge Transfer to In-House Developers
Onboard to project to develop within Release 1
Provide Periodic in-service training on DW techniques
Provide learning materials
Documents
Video
Complete knowledge transfer checklist
Provide Knowledge Transfer to Operational Support
Agile Adoption - Scrum Process
WBS Planning
Operations Management and Troubleshooting

Step 3: Break each deliverable down to individual tasks able to be estimated. If you are using Scrum, these should be small units that can be completed within a Sprint.

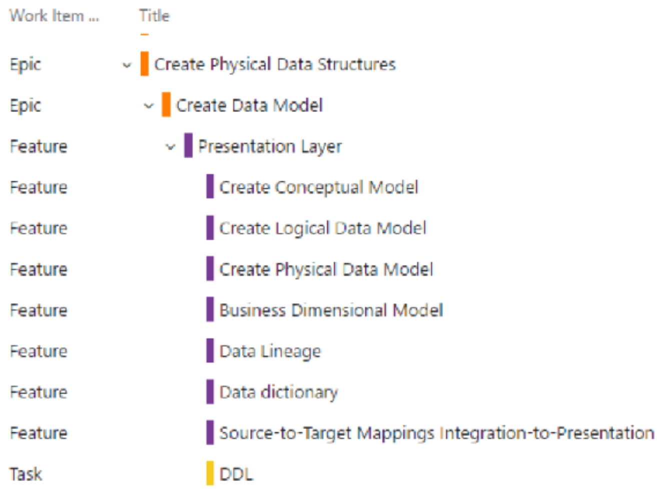


Step 4: Enter the Individual WBS Items in your Scrum Board or Other Tasking Tool

When you get to the lowest level of work products, these become your tasks on a Scrum board. The Scrum board provides many useful capabilities, including measuring work velocity or progress against a backlog.

Every Scrum board tool is different. They are generally hierarchical, reflecting the WBS.

The following illustrates a Scrum board for the Data Model deliverable. Each purple feature drills down into detailed Product Backlog Items or Tasks.

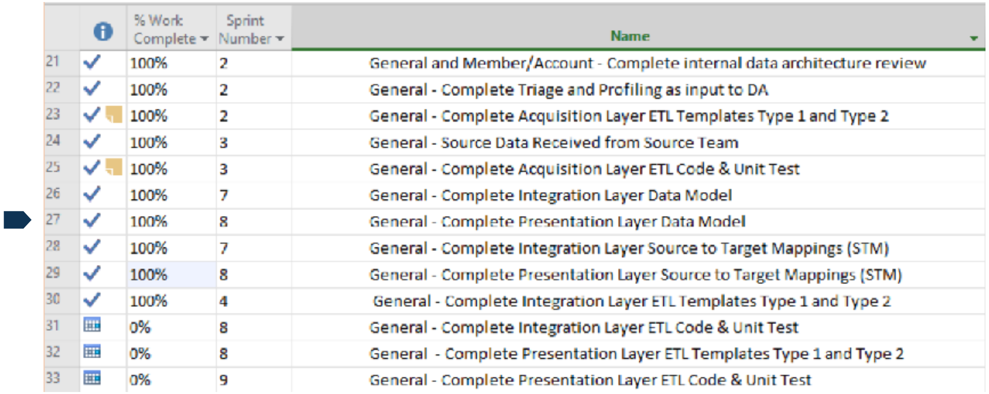


Work Item ...	Title
Epic	Create Physical Data Structures
Epic	Create Data Model
Feature	Presentation Layer
Feature	Create Conceptual Model
Feature	Create Logical Data Model
Feature	Create Physical Data Model
Feature	Business Dimensional Model
Feature	Data Lineage
Feature	Data dictionary
Feature	Source-to-Target Mappings Integration-to-Presentation
Task	DDL

Step 5: Set Milestones

The next step is to determine which tasks will become Milestones, or, checkpoints along the path to completion. Using the example given above, a good Milestone could be stated as a Complete Presentation Layer Data Model. This consists of a number of initial tasks, but provides a certain measure of completion when finalized.

You can see it in Line 27 below.



	ℹ	% Work Complete	Sprint Number	Name
21	✓	100%	2	General and Member/Account - Complete internal data architecture review
22	✓	100%	2	General - Complete Triage and Profiling as input to DA
23	✓ 🗒	100%	2	General - Complete Acquisition Layer ETL Templates Type 1 and Type 2
24	✓	100%	3	General - Source Data Received from Source Team
25	✓ 🗒	100%	3	General - Complete Acquisition Layer ETL Code & Unit Test
26	✓	100%	7	General - Complete Integration Layer Data Model
27	✓	100%	8	General - Complete Presentation Layer Data Model
28	✓	100%	7	General - Complete Integration Layer Source to Target Mappings (STM)
29	✓	100%	8	General - Complete Presentation Layer Source to Target Mappings (STM)
30	✓	100%	4	General - Complete Integration Layer ETL Templates Type 1 and Type 2
31	▦	0%	8	General - Complete Integration Layer ETL Code & Unit Test
32	▦	0%	8	General - Complete Presentation Layer ETL Templates Type 1 and Type 2
33	▦	0%	9	General - Complete Presentation Layer ETL Code & Unit Test

Step 6: Estimate the Lowest Level Tasks

Until this point, task effort has not been estimated. We have only broken down higher level deliverables into measurable components and provided a milestone to indicate when certain checkpoints have been reached.

Every team will have its own way of estimating. It's worth learning how to estimate based on effort, not time, and there are many Scrum resources to help with this. Estimating will happen naturally with collaborative planning to ensure you take all major steps, with tasks broken down into measurable components.

Step 7: Create the Schedule

By breaking down tasks, determining milestones, and completing estimates, you can create a schedule that realistically reflects the necessary work and completion timeline.

PART IV: CONCLUSION

Utilizing a deliverables-based WBS as a project-planning technique provides several benefits:

Creates an effective, repeatable process. If you have WBS from previous projects, you can use it as a starting template for future releases. You will still conduct your collaborative WBS sessions with the team, but they should go much faster since much has already been mapped out.

Evaluates how to break down the tasks and uncover activities and risks that might otherwise have been missed, helping you minimize the negative impact to schedule or scope.

You can identify realistic milestones and estimate and set milestone dates on the schedule, creating a much better idea of whether your project is on track.

[1] http://www.infodocket.com/2014/04/16/how-large-is-the-digital-universe-how-fast-is-it-growing-2014-emc-digital-universe-study-now-available/

[2] Project Management Body of Knowledge (PMBOK®) and the Practice Standard for Work Breakdown Structures – Second Edition, published by Project Management Institute