

Migrating from On-Prem to the Cloud



“The Great Migration” Series

In516ht

/ˈɪnsʌɪt/

definition

noun, plural

A group of people that dig your data,
business and challenges.

"In516ht wins your digital battles."

DWH Development and Modernization

- **15+** years of experience, based in Ljubljana Slovenia
- **60** employees (24 consultants with 10+ years of experience)
- **150+** successfully completed data projects

Development and migrations from traditional on-premises platforms such as IBM, Oracle, Informatica and SAP to modern cloud platforms:

**DEEP INDUSTRY EXPERIENCE
AND BUSINESS KNOW-HOW
WITH FOCUS ON**



BANKING



INSURANCE



TELCO



RETAIL



ENERGY

**+RICH TECHNICAL
EXPERTISE**



PARTNER



What We Do? End-to-End Solutions



Advisory and Consulting

From intuition to fact-based decisions.



Data Management

Ingest, store, organize and maintain data.



Cloud First Strategy

Your Cloud Strategy cookbook.



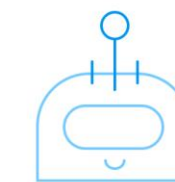
Performance Management

Plan, predict and report.



Managed Services

Give us data and we'll take care of the rest.




Data Science

Infuse next-generation AI and ML capabilities in your business processes

Data Advisory, Mentoring and Trainings

- Most companies already have a data warehouse in place and replacing or migrating it isn't a trivial task of moving data from one database to another.
 - Where should I start?
 - What needs to be discussed?
 - How can I train my people the optimal way?
 - What would the potential cost and benefits be?
- Our senior team is here to help with focused workshops based on discovery meetings to provide guidance and long-term plan of activities needed to become data driven on the Cloud platform.

There is no simple on-premise to cloud migration



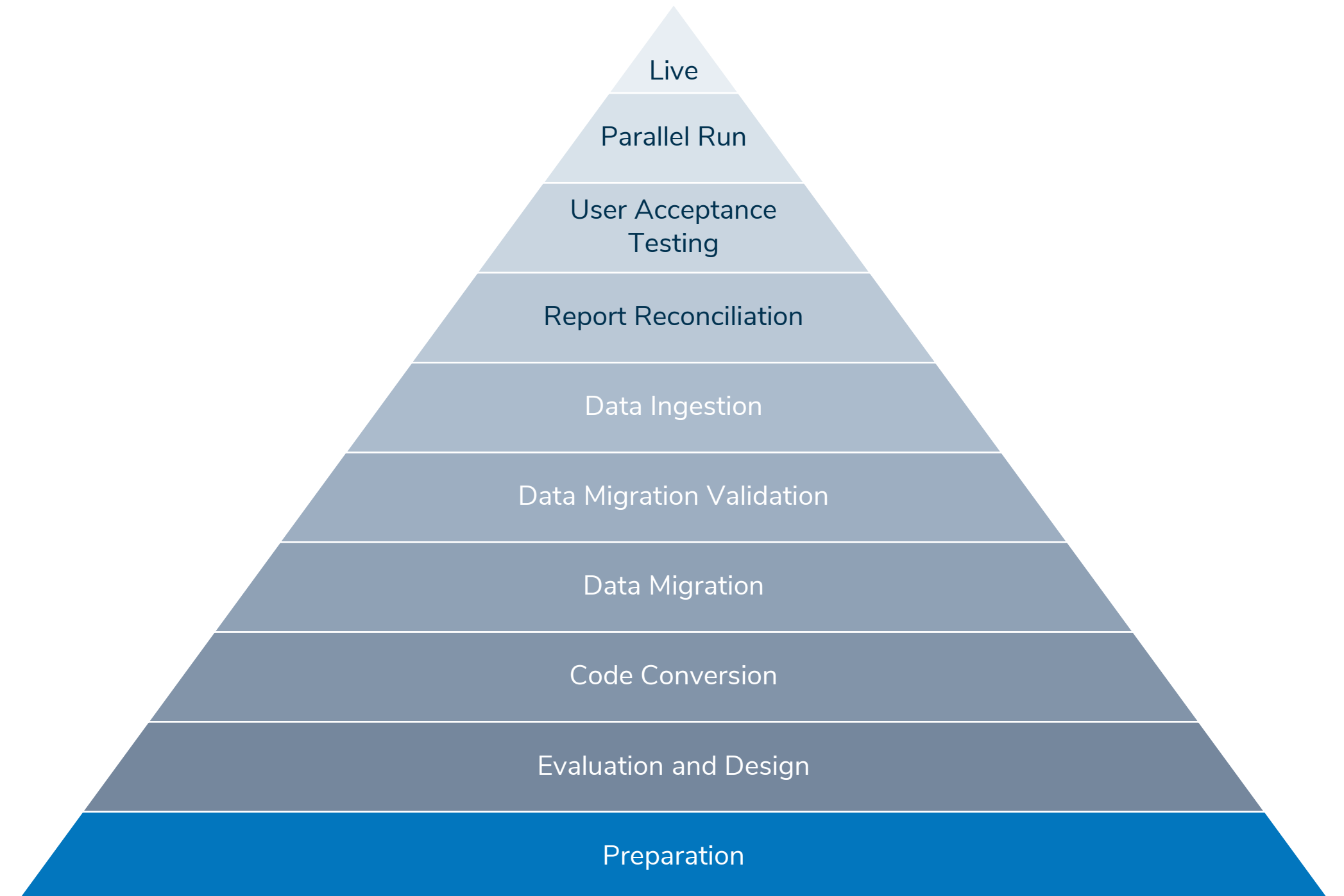
83% of all data migration projects either fail or exceed their budgets and timelines.

Gartner

Preparation is the Key to Success

- Preparation phase:
 - Enablement
 - Migration readiness assessment

Our approach is based on ensuring the least possible risk at every step.



Enablement

- The initial workshop, presentation and brainstorming:
 - To introduce MS Azure Synapse analytics and migration framework
 - To discuss potential first wins
 - To share best practices
 - To debate potential migrations challenges
 - To define brief project overview



Migration Readiness Assessment (MRA)

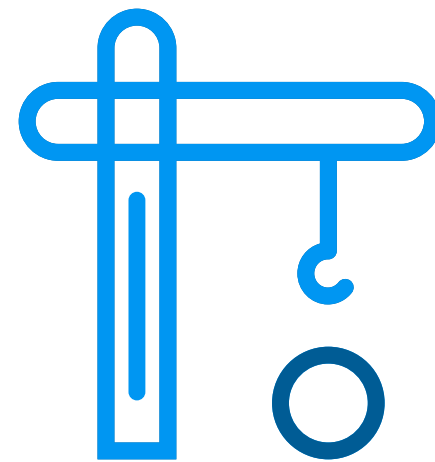
- Focus on workloads, tools, development and deployment processes, migration strategy and project management.
- MRA objectives:
 - To ascertain key information regarding your business and technical needs, expectations and priorities
 - To define and agree on overall architecture, migration strategy, cut-over strategy, system testing & reconciliation strategy
 - To assign project team, project roles and responsibilities
 - To deliver an end-to-end migration schedule, including a high-level plan and detailed task list of migration steps

Evaluation: Scoping

- A successful migration begins with activities to accurately scope the project.
- To be able to define a realistic scope, you need to do a full assessment of your existing data warehouse. For this reason, you should document an as-is architecture and prepare a detailed list of:
 - Databases, including all data warehouse staging areas
 - Data model including all database objects such as tables and views and schemas
 - ETL processes that populate and pull data in a variety of forms
 - Orchestration processes
 - Security schemas and authorization privileges
 - Users and other data consumers
 - BI tools and applications, reports
 - And don't forget about the business logic, data flows, and all the dependencies
- Phased delivery instead of a Big Bang approach.

Migration Strategy

- There is no universal migration strategy.
- The choice of the most suitable approach depends upon your priorities, constraints and primary goals you would like to achieve.



Lift'n' Shift



Lift, Adjust 'n' Shift



Complete Redesign

- Define overall migration approach and subject areas

Proven Project Plan

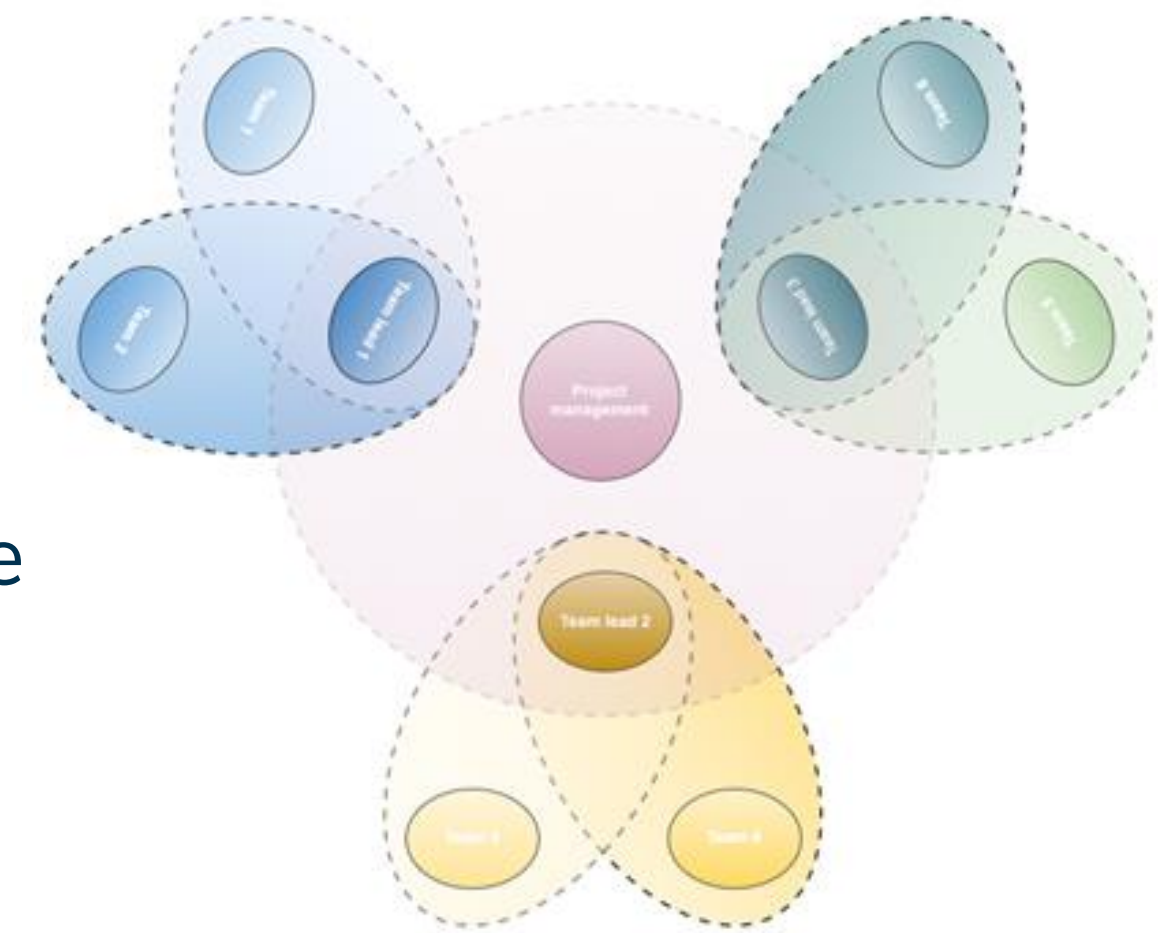
- There is no need to reinvent the processes, we provide a proven high-level plan, so we can focus on:
 - Producing detailed list of tasks
 - Defining milestones to be accomplished in each migration subject area
 - Planning for multiple environments, e.g. DEV, TEST and PROD
- Realistic scope of work can be defined at the end of MRA when migration strategy is clear, effort estimates are done and timeline is defined according to your priorities (project scope, deadlines, resources)
- In case of major unknowns next step is prototyping to estimate the effort

Strong Project Management and Tracking

- We are experienced in various PM methodologies (Prince2, PMP, Scrum) but we always tailor the approach in a way to best serve the customer
- We strive for deliveries at the earliest possible stages of the project making sure that you reap the benefits ASAP
- Project are people. Open and sincere communication is key to ensure that both teams (customer and vendor) are always aligned
 - Clear definition of project roles and responsibilities for all team members
 - Frequent and regular communication supported by weekly and monthly reporting to be able to escalate/fix potential issues in early stage
 - Escalation matrix with clearly defined flow to support quick decisions making
 - Final and intermediate project closures and recognitions

Project Team, Roles and Responsibilities

- Data migration is not only an IT project. It requires as much collaboration from the business side as it does from IT.
- Bigger and more complexity brings more coordination (project management)
- Teams that suits implementation strategy
 - Mixture of customer and our team (training phase, easy knowledge transfer)
 - Near-offshore approach (regular **on-site** workshops and **remote** work organized in virtual teams)
 - **Business** consultant (industry specialist), **technical** consultants (lead architect, mid senior consultants, data engineers) and **project** manager are members of every project



Implementation Phase

- 3 main phases
 - Design (hint: spend enough time on this phase) and detailed business scope
 - Detailed target architecture
 - Setup cloud platform
 - **Detailed project plan for implementation**
 - Development
 - Continuous data feeds
 - Initial load
 - Reconciliation
 - UAT and Parallel Run

Define Target Architecture

- Overall data platform landscape and flow of data in details
 - Public cloud offerings
 - MS Azure Synapse analytics
 - Data ingestion / transformation
 - Reporting, data science platforms and 3rd party integrations
- Exact definition of the data flow
- Which user groups will access which part of the overall data platform

Setup MS Azure Synapse analytics Platform

- One or multiple accounts
- Discussion about DEV, UAT, PROD or any additional environments
- Security setup
- database setup
- Agree naming conventions
- Intermediate areas for reconciliation and tests
- Cloning capabilities
 - for DEV -> UAT and UAT -> PROD and vice versa if needed
 - Business users access cloned schemas

ETL/ELT Process Conversion Analysis

- Scripts
 - ETL code conversion includes all kinds of scripts (DDL, DML, procedures, and other scripts) that must be translated to be valid in MS Azure Synapse analytics
 - Analyse the scripts and estimate the time required for code conversion
- Existing data integration tool
 - Analyse the capabilities to execute jobs as “push-down” to Synapse
 - In Lift & Shift approach the data integration tool must be “near” Synapse
 - Usually, part of replacement phase

Existing data integration is typically the largest unknown in scoping migrations.

Data Integration as a 2-step Approach

- Extraction / ODS phase
 - Extraction phase from the source systems and load into the Synapse (1:1)
 - Different best practices depending on the source system location
 - On-prem much more specific (time-frames, extraction speed, bandwidth)
 - Cloud-to-Cloud lots of solutions and can be “couple of clicks”
 - Extraction phase should be in the new solution as automated as possible (10 or 1000s of tables)
- Transformation phase = the complex part
 - Push-down logic to Synapse is a must
 - “Graphically” transformations (Matillion) or code based (DBT, DataMerlin,...)

Initial Data Migration

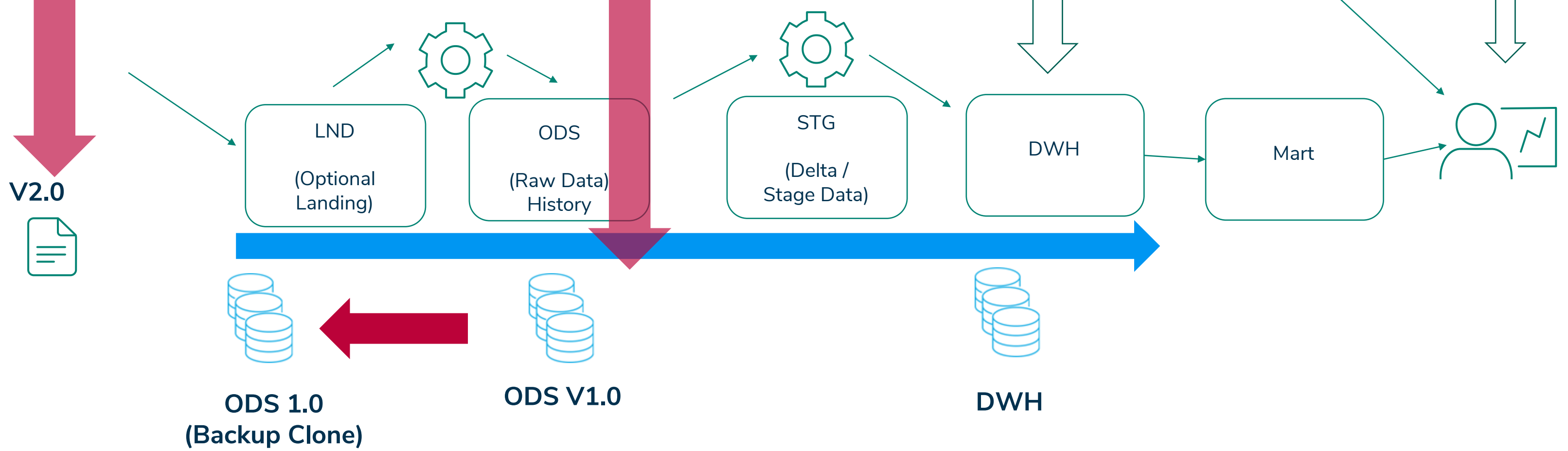
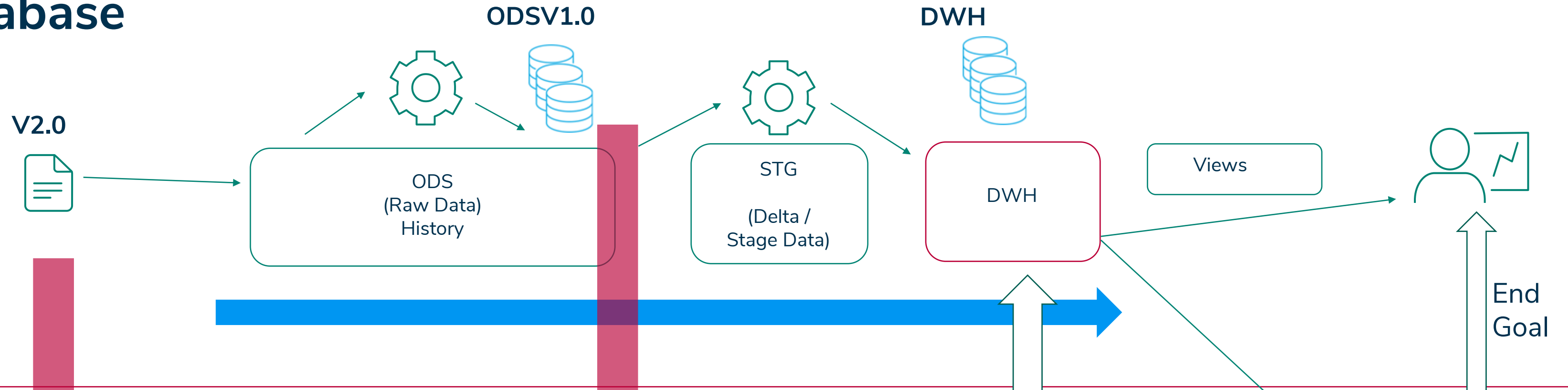
- The movement of an initial dataset from the legacy system to Synapse.
- Determine list and size of tables from existing DWH systems
- Important elements to focus on
 - Extraction speed from the existing DWH environment
 - Available bandwidth to move the data to Synapse
- How to extract deltas from the existing DWH system?

Data Validation / Reconciliation

- Most crucial step to automate as much as possible.
- Establish data validation expectations (row count, hash, summary of data).
- Compare data results in the development phase (use cloning to have DevOps capabilities).
- Build data quality checks / tests that run automatically on each load.
- Determine remediation plan for data validation issues.

Test Example

Legacy Database



User Acceptance Testing & Parallel Run

- In majority of use cases the business will do the UAT using the BI tool – comparing side by side reports.
- Any feedback / change is a great candidate for an additional data quality process to be implemented in the regular DWH loading.
- Parallel Run
 - Almost mandatory if technically possible
 - If possible, implement daily / weekly delta extract from the existing DWH and load it to Synapse to do additional automatic tests
 - Additional confidence for the client that the migration was done successfully
 - Can last from 1-3 months from our experience

Tools to Accelerate the Migration Process

- Migration is an iterative process. Having as much automation as possible minimizes project risk and brings many invaluable advantages.
 - Automated code conversion cuts down migration time by factors
 - Automated technical testing assures that the data is properly migrated
 - Meta-data driven approach to data integration brings ability to repeat data transformation and apply new changes with almost no effort

Our assets can make your project much faster, more efficient and far less risky in comparison to traditional methods.

Code converter
the code conversion toolkit

antFarm
the lift 'n' shift tool

DataMerlin
the DWH automation platform





More Information

www.in516ht.com

sales@in516ht.com

klemen.logar@in516ht.com

in516HT
know your numbers