Northwind BI solution

Project Scope

|  |  |  |
| --- | --- | --- |
| High business value |  | * See sales with and without discounts. * See average discount by customers and geography. * See sales on holidays, weekdays and weekends. * See sales by salespersons and employee geography. |
| Low business value |  |  |
|  | Low feasibility | High feasibility |

# Required BI Software

|  |  |
| --- | --- |
| Software | Rationale |
| SQL Server Database Engine | * The CEO wants a consistent view of all business data, and a centralized data warehouse in SQL Server would provide this. |
| SQL Server Integration Services | * The business data required for analysis and reporting is currently spread across a range of data sources. Integration Services will provide an ETL platform to populate and refresh the data warehouse. |
| SQL Server Master Data Services | * The CEO has complained about inconsistent data, which could potentially be caused by a lack of central data management for key business entities. |
| SQL Server Data Quality Services | * The specialists have difficulty ensuring that their analysis of sales data is accurate because of quality issues in the data. Data Quality Services could be used to cleanse records and improve the consistency, thereby reducing the inaccuracy of the sales analysis. |
| Power BI Report Server | * The database administrators team needs to receive email notifications about errors in the ETL process. * Data stewards need to be notified of errors and raw values in data cleansing processes. |

# BI Topology

## Single BI server



* SQL Server DB Engine
  + Data Warehouse
  + Staging database
  + SSIS catalog
  + Reporting Services catalog
  + DQS catalog
  + MDS catalog
* SQL Server Integration Services
* Power BI Report Server
* Data Quality Services
* Master Data Services

SRVBI



CLIENT PC

* Microsoft Excel
  + Power Query
  + Power Pivot
  + Power View
* Report Builder
* Power BI Desktop
* Microsoft Edge

Notes:

* Minimal server hardware and software license requirements, but the server would require significant memory, CPU, and disk resources for all but the most lightweight BI workloads.
* The range of different workload types on the server would make it difficult to specify and configure hardware resources appropriately.
* The server could be clustered to provide high availability.

## Dedicated report server



* SQL Server DB Engine
  + Data Warehouse
  + Staging database
  + SSIS catalog
  + DQS catalog
  + MDS catalog
* SQL Server Integration Services
* Data Quality Services
* Master Data Services

SRVDWH



CLIENT PC

* Microsoft Excel
  + Power Query
  + Power Pivot
  + Power View
* Report Builder
* Power BI Desktop
* Microsoft Edge



* SQL Server DB Engine
  + Reporting Services catalog
* Power BI Report Server

SRVPBIRS

Notes:

* The reporting and DWH workloads are full separated.
* Using a dedicated data warehouse server makes it easier to manage the ETL and data warehouse workloads.
* Additional report servers could be added to scale out reporting.
* One or both servers could be clustered to provide high availability.

## Dedicated ETL server



* SQL Server DB Engine
  + Data Warehouse
  + Staging database

SRVDWH



CLIENT PC

* Microsoft Excel
  + Power Query
  + Power Pivot
  + Power View
* Report Builder
* Power BI Desktop
* Microsoft Edge



* SQL Server DB Engine
  + Reporting Services catalog
* Power BI Report Server

SRVPBIRS



* SQL Server DB Engine
  + SSIS catalog
  + DQS catalog
  + MDS catalog
* SQL Server Integration Services
* Data Quality Services
* Master Data Services

SRVETL

Notes:

* The DWH, reporting and ETL workloads are full separated.
* The staging database could be hosted on the ETL server or data warehouse server depending on workloads.
* SRVDWH, SRVPBIRS, and SRVETL could be clustered to provide high availability.
* Additional report servers and ETL could be added to scale out.

## Dedicated master data server



* SQL Server DB Engine
  + Data Warehouse
  + Staging database

SRVDWH



CLIENT PC

* Microsoft Excel
  + Power Query
  + Power Pivot
  + Power View
* Report Builder
* Power BI Desktop
* Microsoft Edge



* SQL Server DB Engine
  + Reporting Services catalog
* Power BI Report Server

SRVPBIRS



* SQL Server DB Engine
  + DQS catalog
  + MDS catalog
* Data Quality Services
* Master Data Services

SRVMDM



* SQL Server DB Engine
  + SSIS catalog
* SQL Server Integration Services

SRVETL

Notes:

* The DWH, reporting ETL and MDM workloads are full separated.
* The staging database could be hosted on the ETL server or data warehouse server depending on workloads.
* Servers could be clustered to provide high availability.
* Additional report servers and ETL could be added to scale out.

## Distributed Report Server



* SQL Server DB Engine
  + Data Warehouse
  + Staging database

SRVDWH



CLIENT PC

* Microsoft Excel
  + Power Query
  + Power Pivot
  + Power View
* Report Builder
* Power BI Desktop
* Microsoft Edge



* Power BI Report Server

SRVPBIRS



* SQL Server DB Engine
  + DQS catalog
  + MDS catalog
* Data Quality Services
* Master Data Services

SRVMDM



* SQL Server DB Engine
  + SSIS catalog
* SQL Server Integration Services

SRVETL



* SQL Server DB Engine
  + Reporting Services catalog

SRVPBIRSDB

Notes:

* The DWH, reporting ETL and MDM workloads are full separated.
* The staging database could be hosted on the ETL server or data warehouse server depending on workloads.
* The report server uses a separate database server to host the report catalog.
* Servers could be clustered to provide high availability.
* Additional report servers and ETL could be added to scale out.

# Расчёт параметров оборудования для хранилища данных

Перед началом проектирования системы хранилища данных необходимо рассчитать maximum consumption rate (MCR) для одного ядра процессора в сочетании с ядром базы данных SQL Server.



# Проектирование хранилища данных

## Проектирование логической схемы ХД

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Бизнес-процессы | Измерения | | | | | | |
| **Время** | **Сотрудник** | **Оптовые покупатели** | **Розничные покупатели** | **Продукты** | **Регионы продаж** | **Маркетинговые Акции** |
| Оптовые продажи | Х | X | Х |  | Х | Х |  |
| Розничные продажи | Х |  |  | Х | Х |  | Х |

**Примечания:**

* Географические данные (страна, регион, город) используются в нескольких измерениях, так как при анализе продаж географические данные могут быть важны рекомендуется вынести в отдельное измерение.

### Модель данных бизнес-процесса «Оптовые покупатели»

Время

(Заказа, Доставки)

Год

Квартал

Месяц

День

Оптовый покупатель

Сотрудник

Товар

Регион продаж

Категория

Подкатегория

Товар

Регион продаж

Страна продаж

Территория продаж

Менеджер

Продавец

Тип

Магазин

Страна

Провинция

Город

Магазин

### Модель данных бизнес-процесса «Розничные покупатели»

Время

(Заказа, Доставки)

Год

Квартал

Месяц

День

Розничный покупатель

Маркетинговая акция

Товар

Категория

Подкатегория

Товар

Тип

Маркетинговая

акция

Пол

Материальный

статус

Расстояние до

магазина

Количество машин

Страна

Провинция

Город

Магазин

## Проектирование физической реализации ХД

# Проектирование ETL

# Проектирование аналитических моделей данных

# Проектирование решения для доставки бизнес-аналитики

# Мониторинг и оптимизация решения бизнес-аналитики

# Управление решением бизнес-аналитики