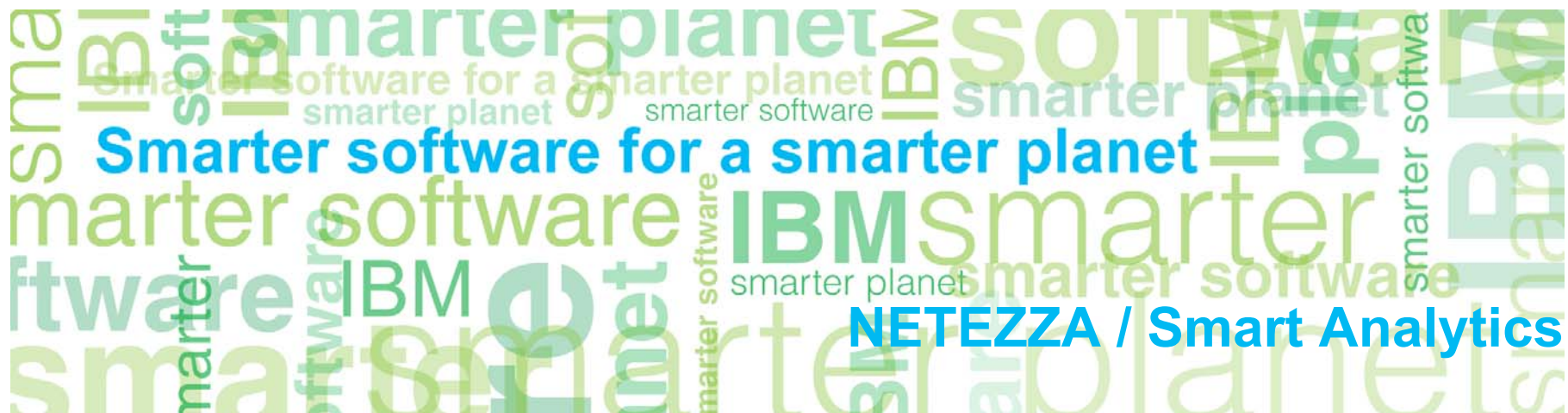


# Gyors adattárház működés? Van megoldás!

Optimalizált adattárház rendszerek: Netezza, Smart Analytics System



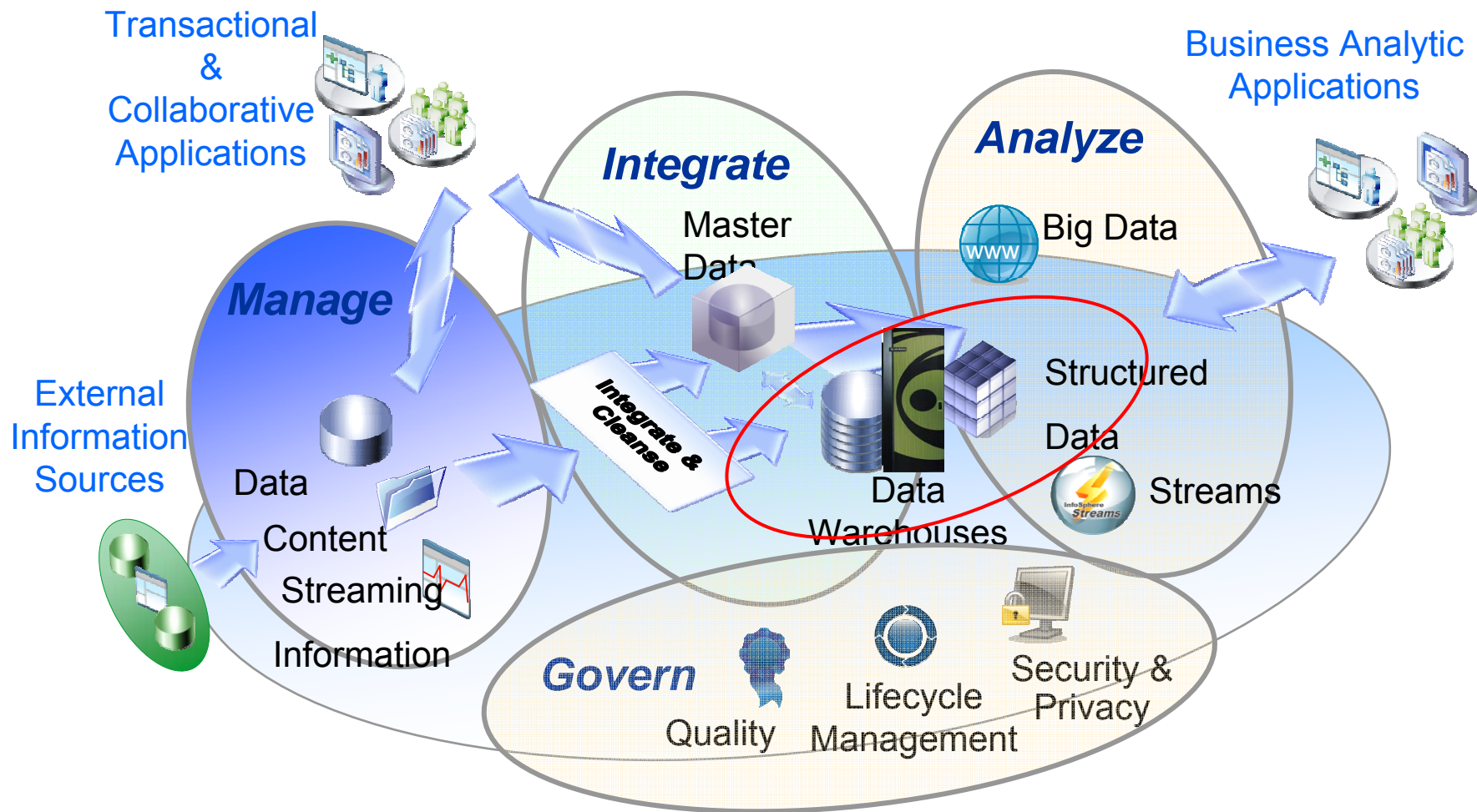
## Napirend

- Adattárház helye az informatikai rendszerben
  - IBM Warehouse megoldások: Teljeskörű integráltság
  - „Egyszerűségtől az egyedi megoldásokig”
  - Tradicionális adattárház
  
- Optimalizált adattárház rendszerek
  - Netezza adattárház célszámítógép / célhardver
    - Mikor jó a Netezza
    - „Célhardver egyszerűsége”
    - IBM Netezza TwinFin architektúra
    - Netezza skálázhatóság
  
  - Flexibilis adattárház és analitika: Smart Analytics System
    - Hatékony adattárház
    - Moduláris felépítés
    - Nagy rendelkezésre állás, nagyvállalti megoldások
  
- Kérdések, válaszok

## Napirend

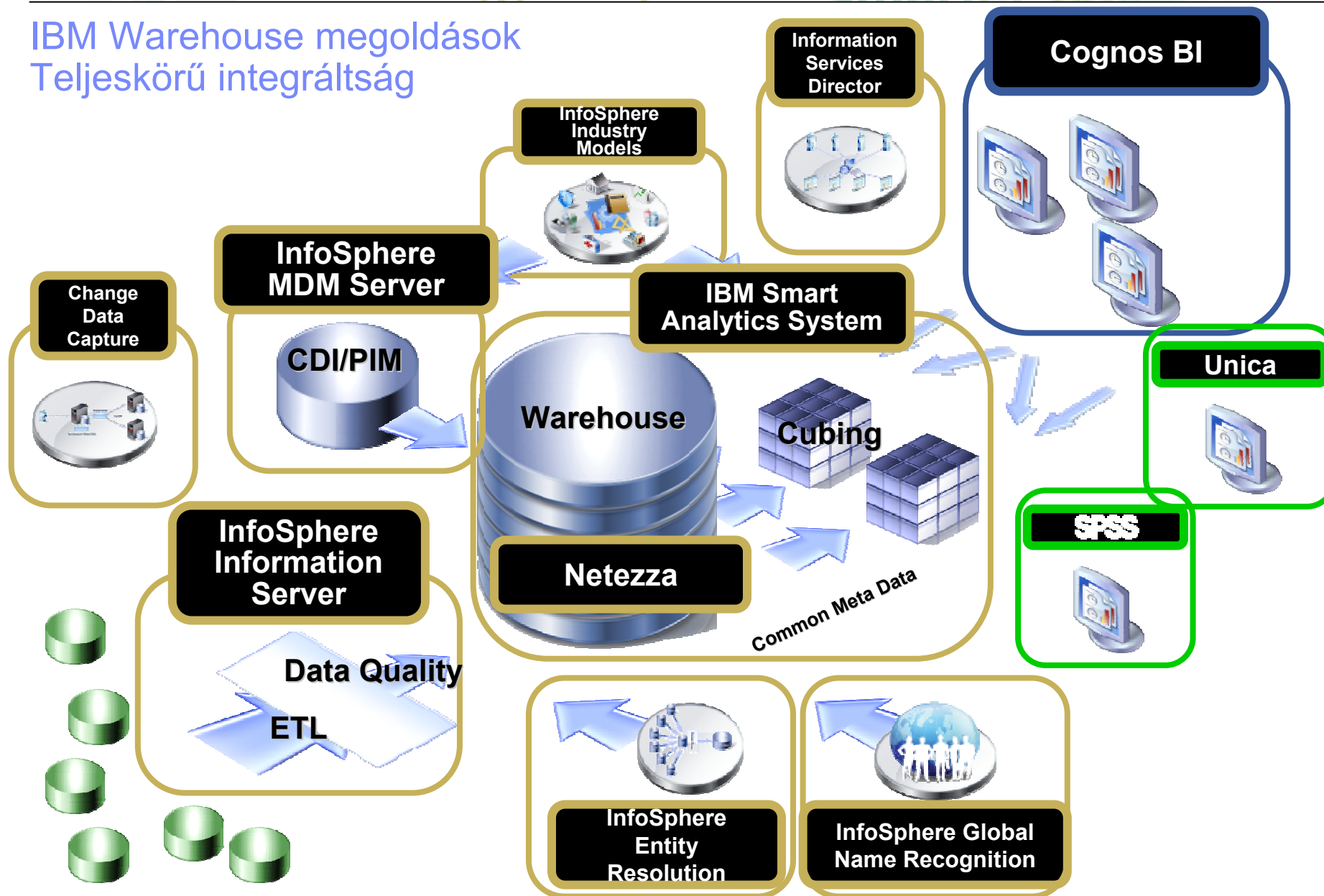
- Adattárház helye az informatikai rendszerben
  - IBM Warehouse megoldások: Teljeskörű integráltság
  - „Egyszerűségtől az egyedi megoldásokig”
  - Tradicionális adattárház
  
- Optimalizált adattárház rendszerek
  - Netezza adattárház célszámítógép / célhardver
    - Mikor jó a Netezza
    - „Célhardver egyszerűsége”
    - IBM Netezza TwinFin architektúra
    - Netezza skálázhatóság
  
  - Flexibilis adattárház és analitika: Smart Analytics System
    - Hatékony adattárház
    - Moduláris felépítés
    - Nagy rendelkezésre állás, nagyvállalti megoldások
  
- Kérdések, válaszok

# Adattárház helye az informatikai rendszerben





# IBM Warehouse megoldások Teljeskörű integráltság



## Egyszerűségtől az egyedi megoldásokig

**Netezza**



**Célhardver**

**Smart Analytics System InfoSphere Warehouse**



**Kiegészítő / Integrált**

**InfoSphere Warehouse**

**Informix Warehouse**

**Informix Warehouse Accelerator**



**Egyedi megoldás**

**Warehouse Accelerators**

**Information Management Portfolio**  
*(Information Server, MDM, Streams, etc)*

**Egyszerűség**

**Egyediség**

Egyszerűség, rugalmasság, a választás szabadsága

## IBM Data Warehouse & Analytics Solutions

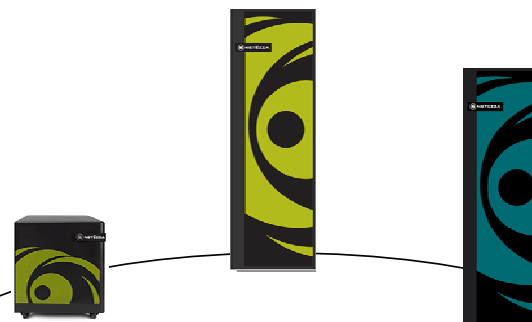
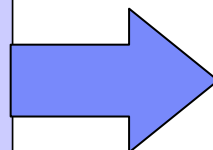
IBM Netezza	IBM Smart Analytics System	IBM InfoSphere Warehouse
<p><b>Célhardver</b></p> <ul style="list-style-type: none"> <li>• Speciális adattárházra, ad-hoc analízisre optimalizált, 10x – 100X sebesség</li> <li>• Előtérben az analitikai lekérdezések</li> <li>• Az eszközök nem testreszabhatók, fizikai skálázhatóság</li> <li>• IBM, előkészített megoldás</li> <li>• Minimális adminisztráció, DBA, ügyfélerőforrás</li> </ul>	<p><b>Integrált rugalmas</b></p> <ul style="list-style-type: none"> <li>• IBM által összehangolt szoftver és hardver csomag (fejlett DB2, Cognos funkcionalitás)</li> <li>• Komponenspaletta</li> <li>• Rugalmas megoldás különböző adat, felhasználói modulok kezelésére (egyaránt használható OLTP-re, DW-re)</li> <li>• Fizikai és logikai skálázhatóság (modulok, hangolás)</li> <li>• Ár/teljesítmény „jóméretesíthető”</li> </ul>	<p><b>Egyedi megoldás</b></p> <ul style="list-style-type: none"> <li>• Ügyféllel közösen kialakított egyedi megoldás</li> <li>• Szerverek, tároló eszközök, adatbázisok és analitikai szoftverek széles választéka</li> <li>• Mind IBM, mind nem IBM megoldások egyben</li> </ul>



## IBM Warehousing & Analytics – megoldások pozicionálása

### Elsődleges végfelhasználói elvárások

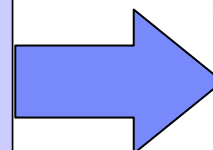
1. Nagy teljesítményű analitikai lekérdezések külső beavatkozás, hangolás nélkül
2. Nincs tárolóeszköz adminisztráció
3. Leggyorsabb bevezetés



**IBM Netezza**  
(Appliances)

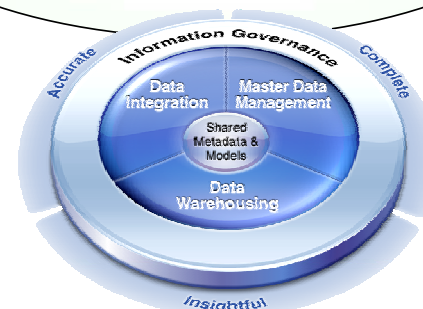
### Elsődleges végfelhasználói elvárások

1. Nagy teljesítményű analitikai lekérdezések és on-line tranzakciók együttesen fontosak
2. Power Systems és System Z platform
3. Egységes, integrált DB2 használat

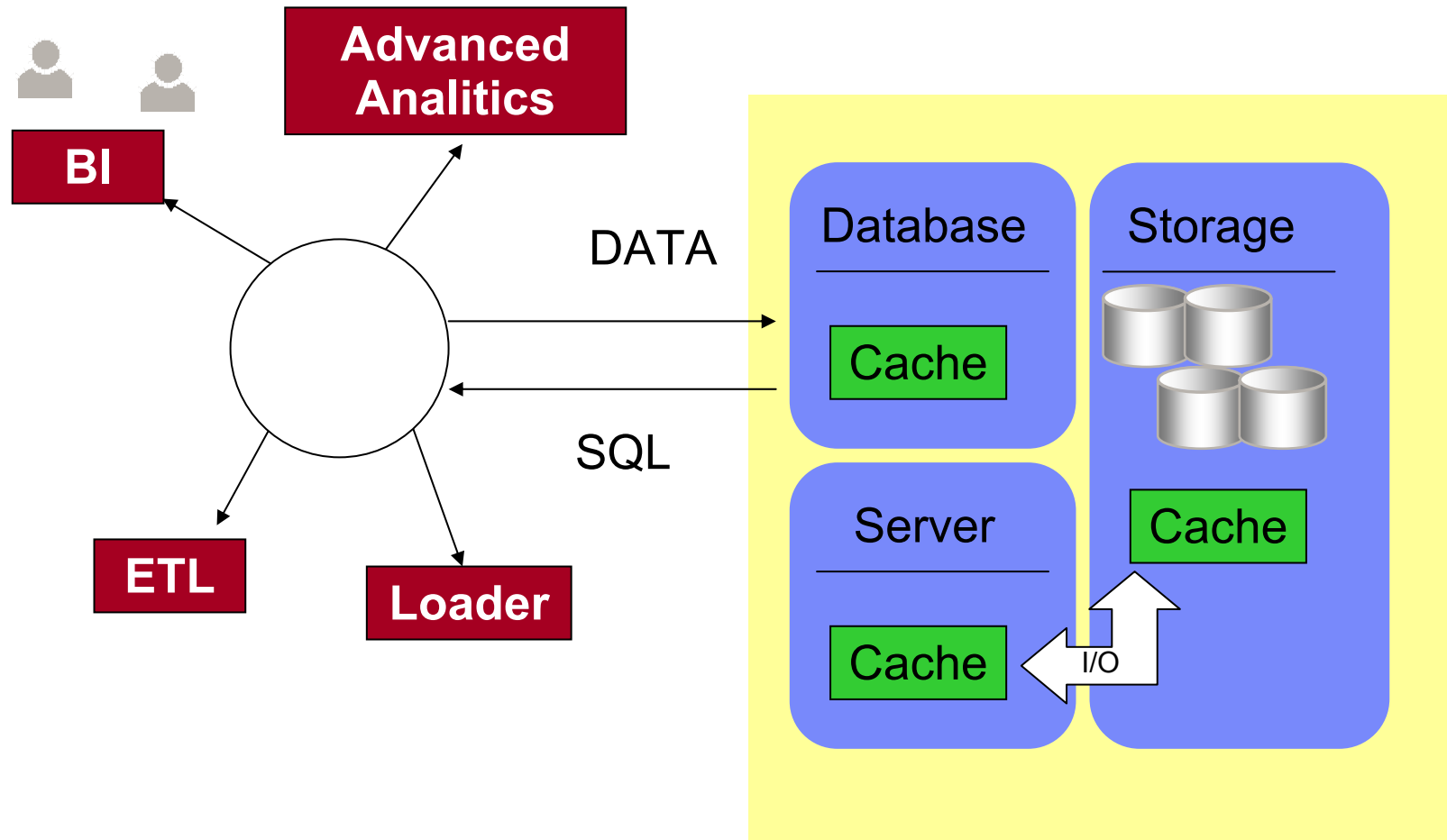


**IBM Smart Analytic System**  
(Optimized systems)

**IBM InfoSphere Warehouse**  
(Custom configurations)



## Tradicionális adattárház

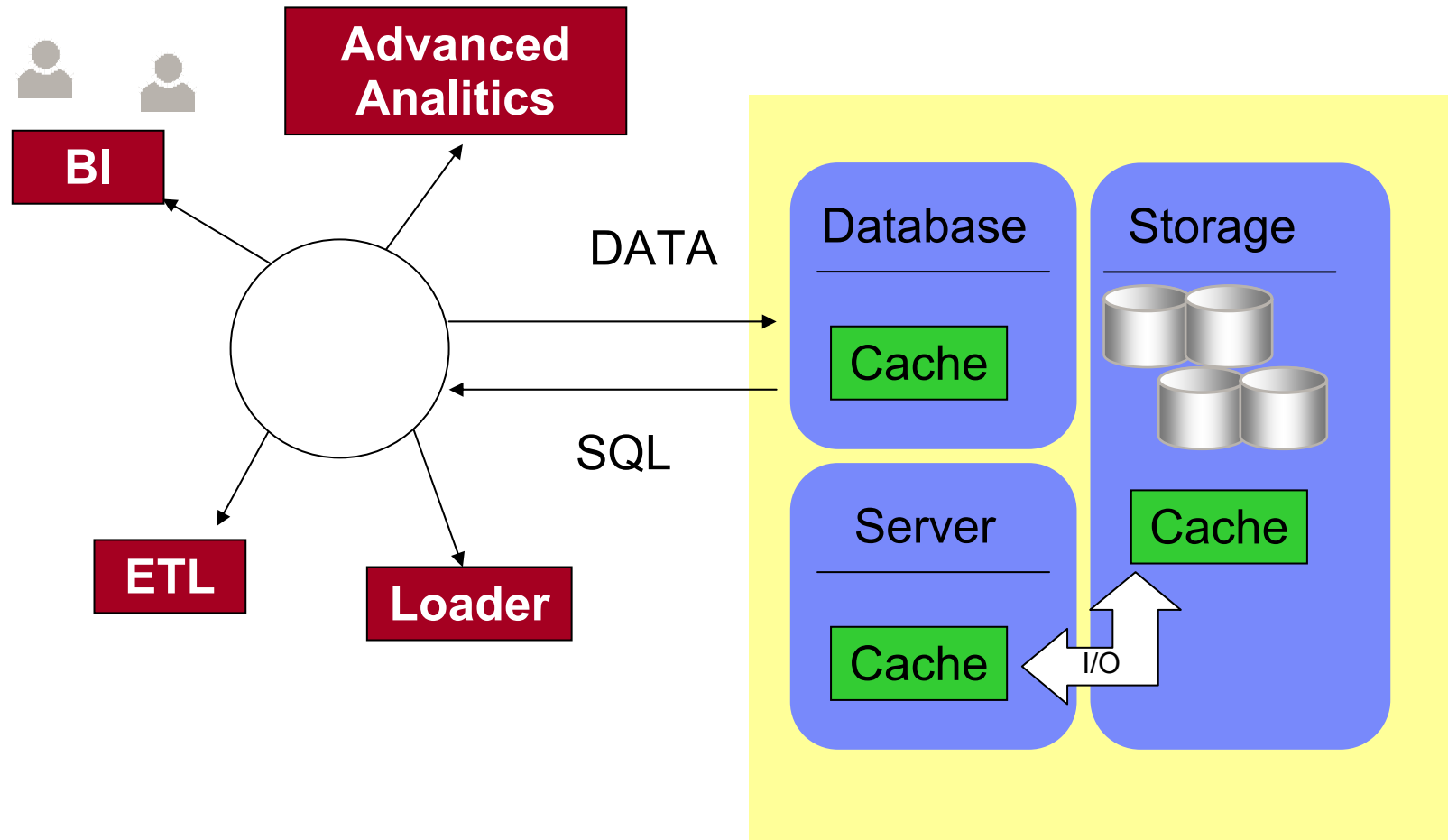




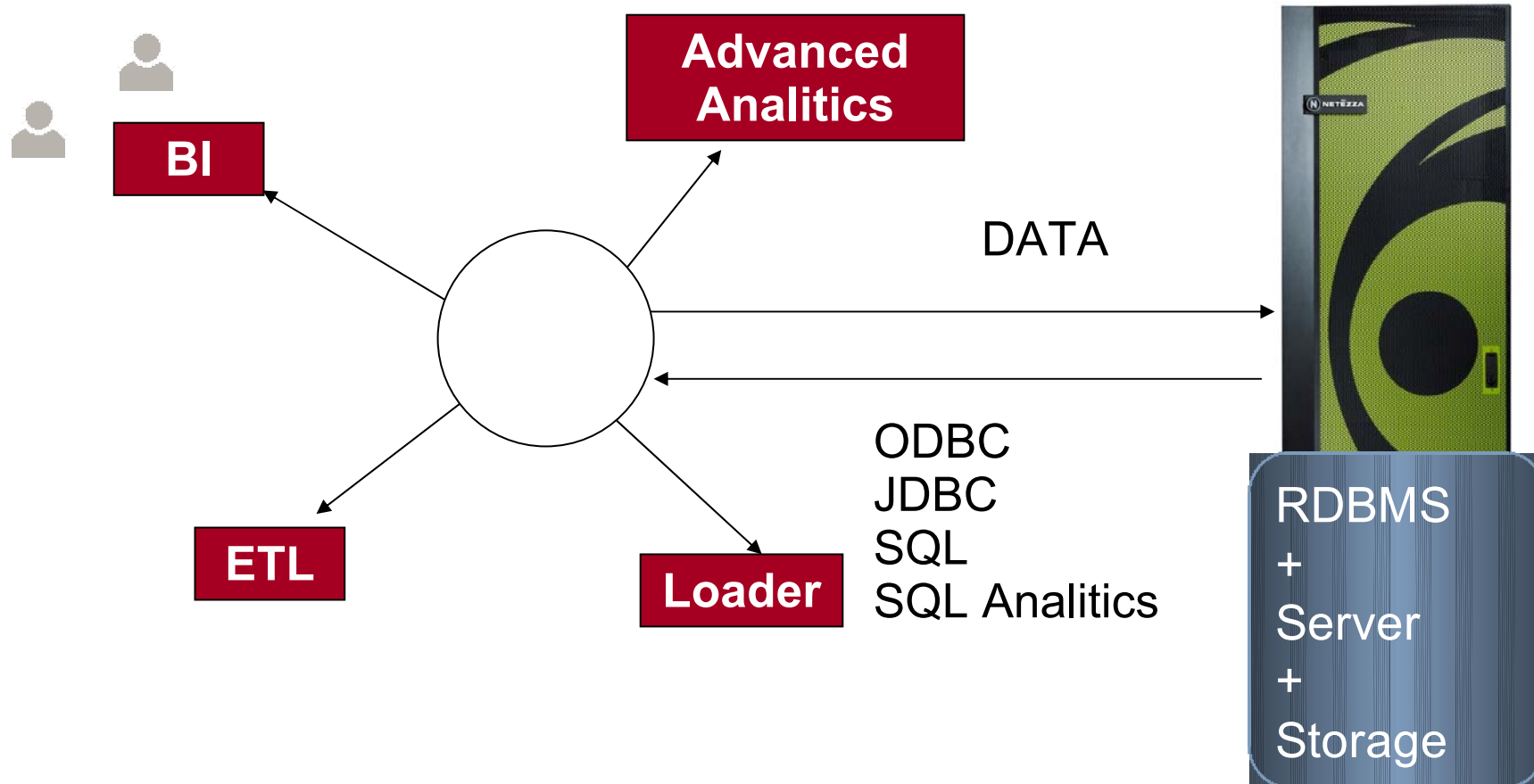
## Napirend

- Adattárház helye az informatikai rendszerben
  - IBM Warehouse megoldások: Teljeskörű integráltság
  - „Egyszerűségtől az egyedi megoldásokig”
  - Tradicionális adattárház
  
- Optimalizált adattárház rendszerek
  - Netezza adattárház célszámítógép / célhardver
    - Mikor jó a Netezza
    - „Célhardver egyszerűsége”
    - IBM Netezza TwinFin architektúra
    - Netezza skálázhatóság
  
  - Flexibilis adattárház és analitika: Smart Analytics System
    - Hatékony adattárház
    - Moduláris felépítés
    - Nagy rendelkezésre állás, nagyvállalti megoldások
  
- Kérdések, válaszok

## Tradicionális adattárház



## Netezza adattárház célszámítógép / célhardver



## Mikor jó a Netezza

- Nagy mennyiségű adat, ideális: 5TB - 20TB
- Folyamatos nagymértékben növekvő adatmennyiség (fontos a jó skálázhatóság)
- Egyszerűség kulcskérdés
- Elemzések túl hosszú idejűek, hangolás már költséges
- Valódi komplex BI analízis sok ad-hoc lekérdezések (query)
- Rövid bevezetési idő (time-to-value)
- Komplex analitika gyorsítása: Netezza I-class
- Meglévő rendszerekkel jól integrálható adattárház

## TwinFin™ célhardver



- **Sebesség**: 10-100x mint tradicionális tárház.
  - AMPP: speciális, masszívan párhuzamos architektúra (FPGA alapú hardvergyorsítás)
- **Skálázhatóság**: 1T – 1,5 Pb
- **Smart**: analitikai funkciók gyorsítása
  
- „Célhardver egyszerűsége”
  - kevesebb adminisztráció
  - kevesebb hangolás
- Kiforrott technológia sok ügyféllel
- Olcsóbb mint hasonló ugyanilyen teljesítményű rendszerek

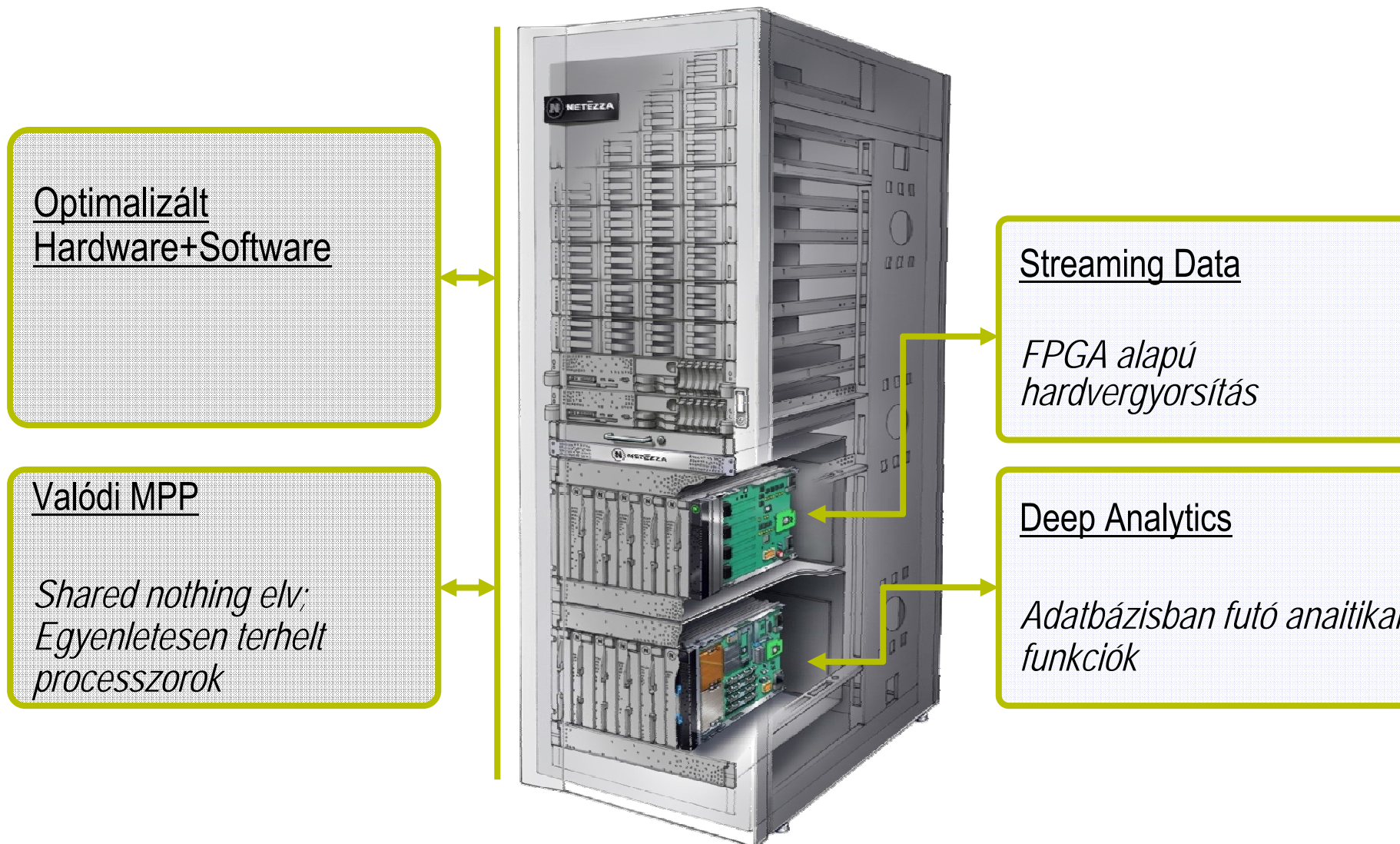


## Egyszerű telepítés és üzemeltetés

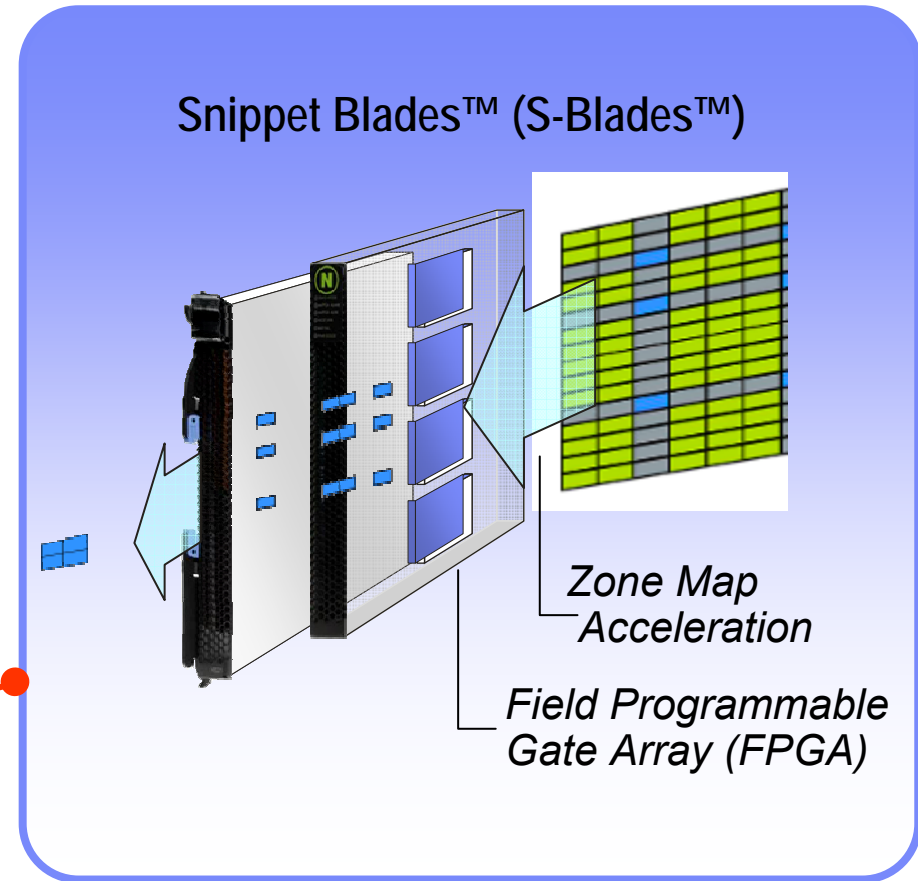
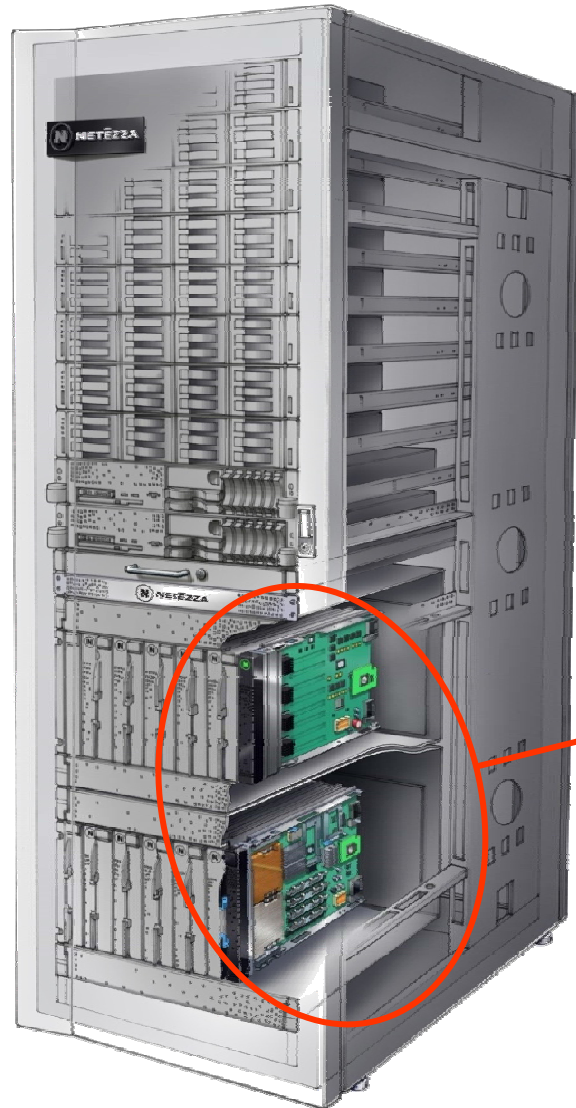
- **Üzemeltetés**
  - töltsd fel adattal és használd
  - installáció (time-to-value) ~ napok
- **BI fejlesztők**
  - nincs hangolási szükséglet, nem kell konfigurálni indexeket menedzselni
    - out-of-the-box performance
- **ETL fejlesztők**
  - gyors betöltés és adattranszformáció
    - gyors ETL logika & in-database transformation
- **Üzleti felhasználók**
  - kisebb válaszidők
    - betöltés, lekérdezés párhuzamosan
  - valódi ad-hoc analitika gyorsítása



## IBM Netezza TwinFin™ architektúra

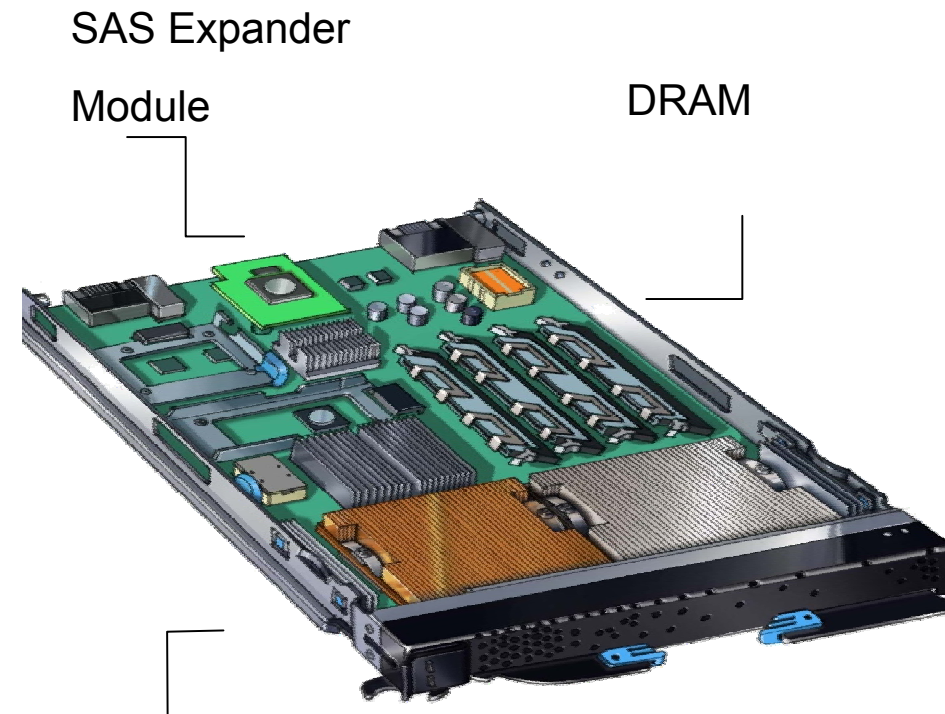


## Netezza S-BLADE megoldás

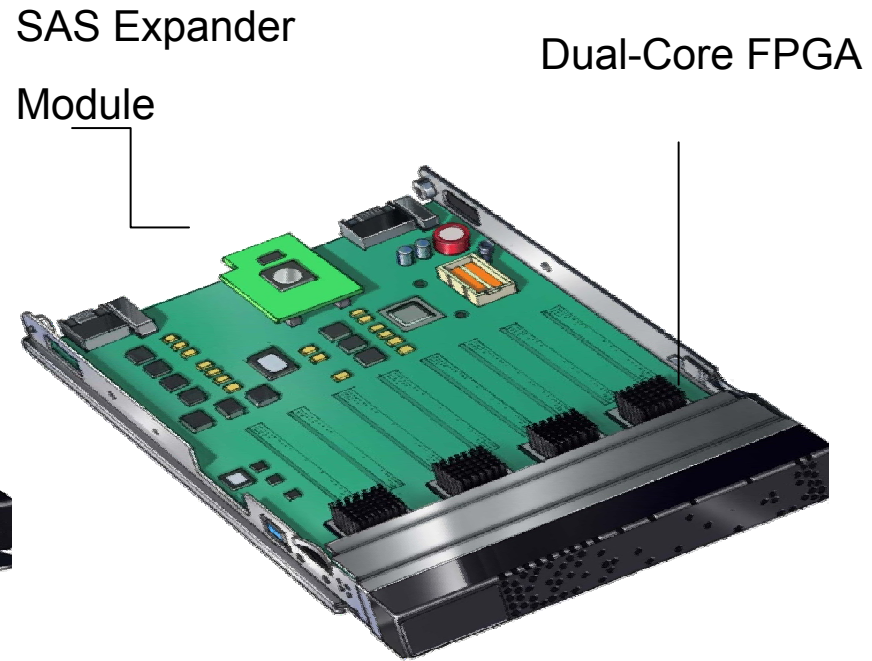


- Hardver alapú adatfolyam feldolgozás
- Lekérdezés szűrés közel az adattároláshoz

## S-Blade™ felépítése



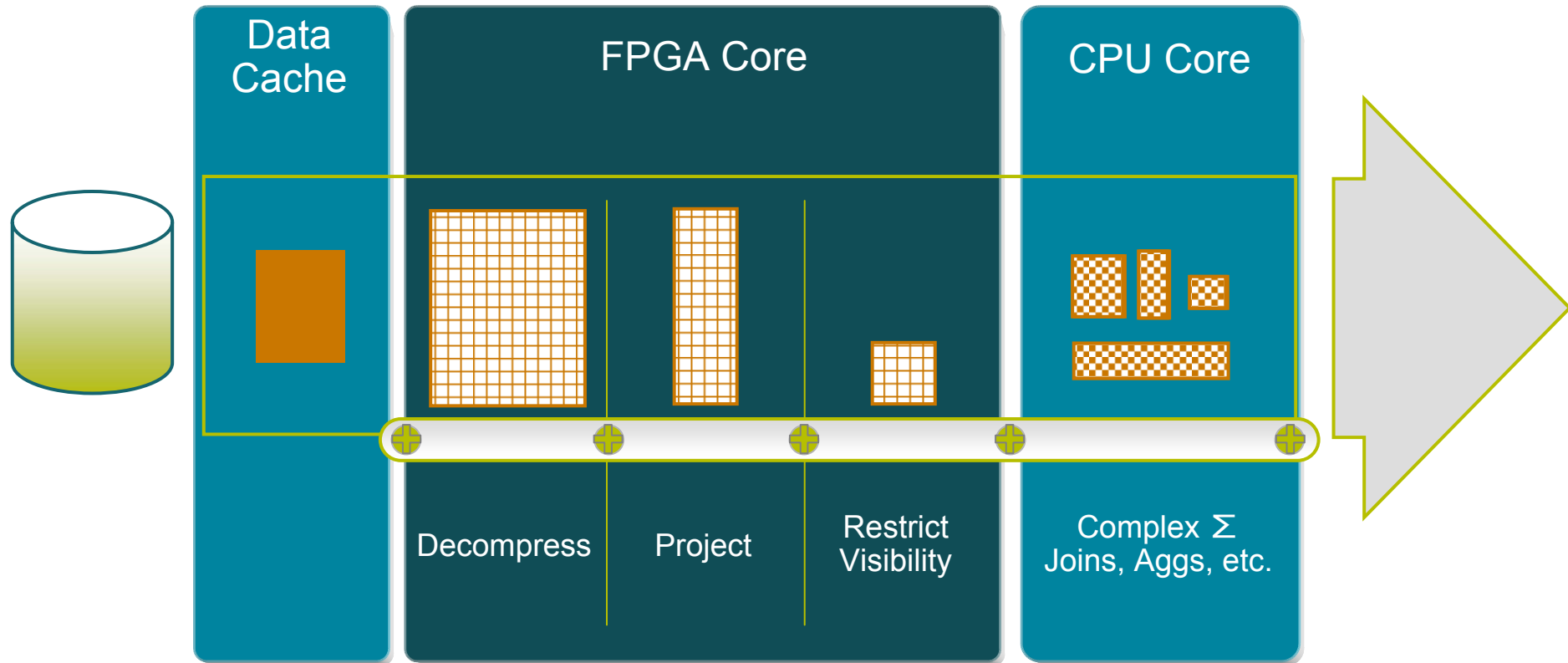
IBM BladeCenter Server



Netezza DB Accelerator



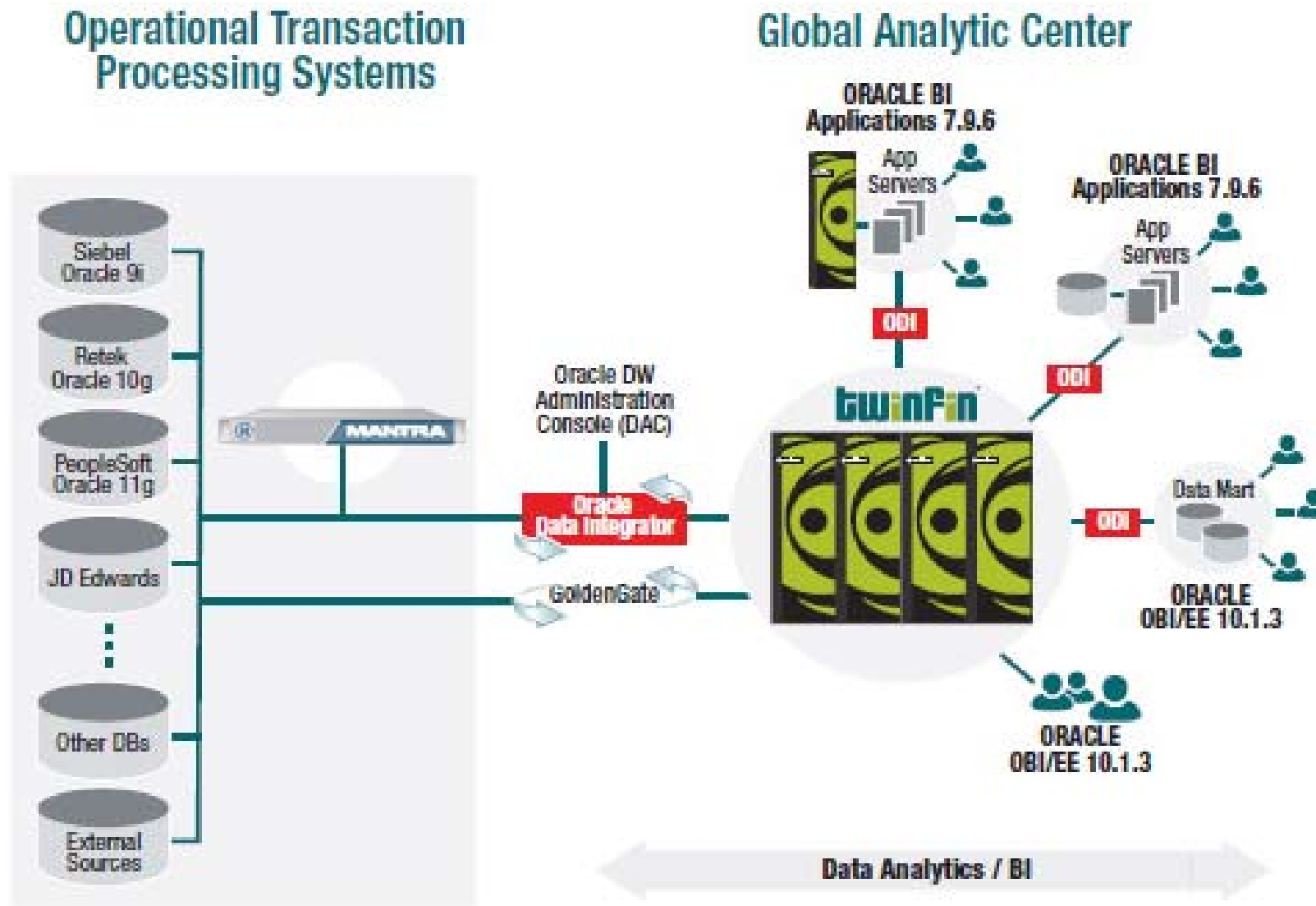
## S-Blade adatfolyam



```
Select sex, age, count(*)  
From MultiBillionRowTable  
Where BirthDate < '01/01/1967'  
And PostCode like 'SW%'  
Group by sex, age;
```



## Netezza – Oracle rendszerintegráció



## Netezza integráció – BETÖLTÉS

### Data Integration

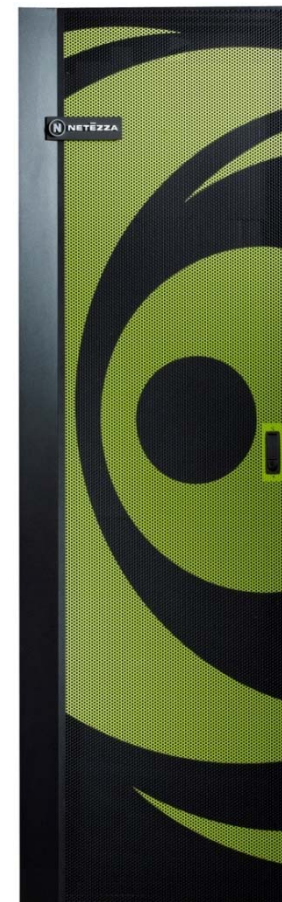
Integrátor eszközök:

- Business Objects/SAP
- GoldenGate Software (Oracle)
- Informatica
- IBM Information Server
- Sunopsis (Oracle)

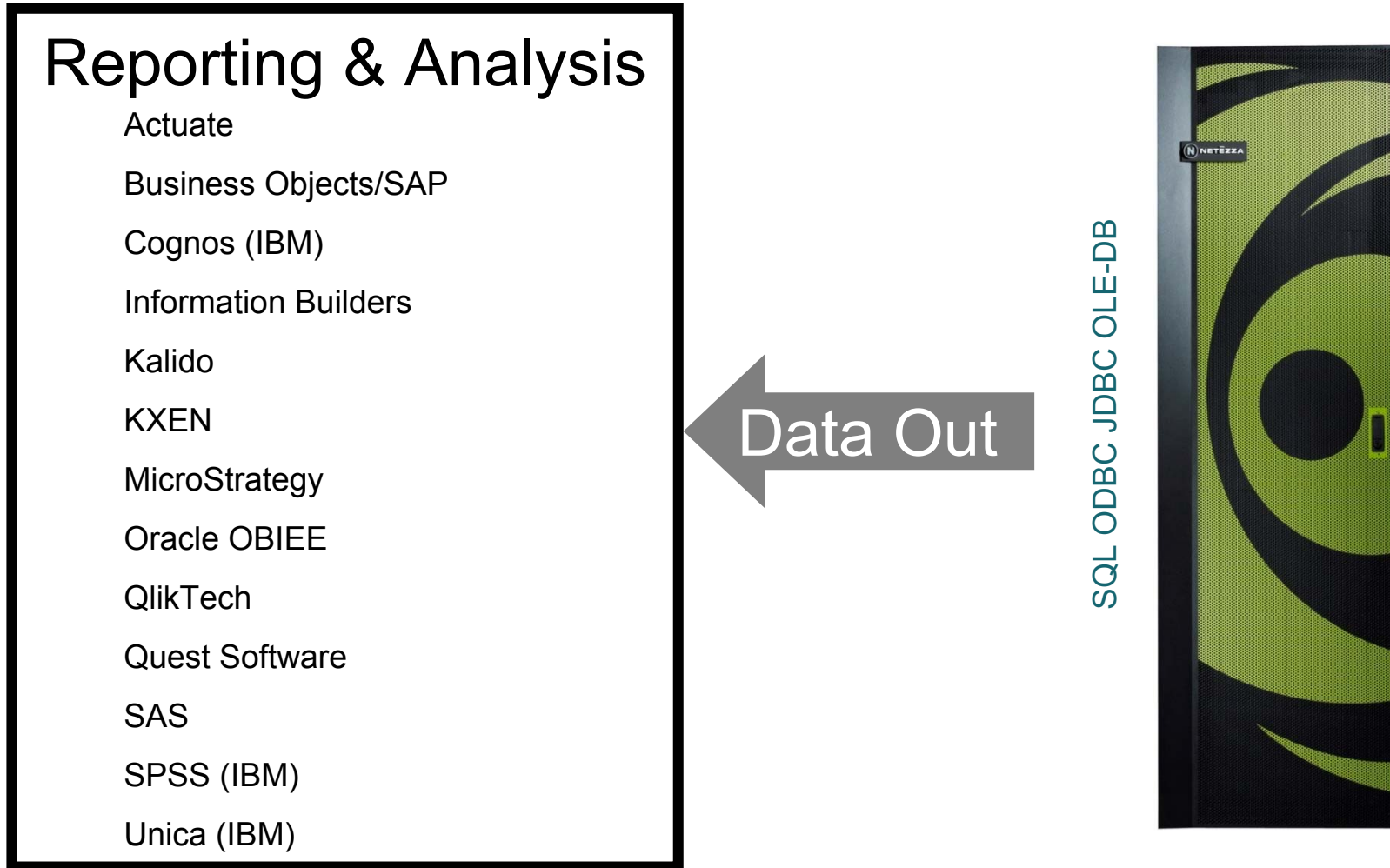
Constant 2TB /óra



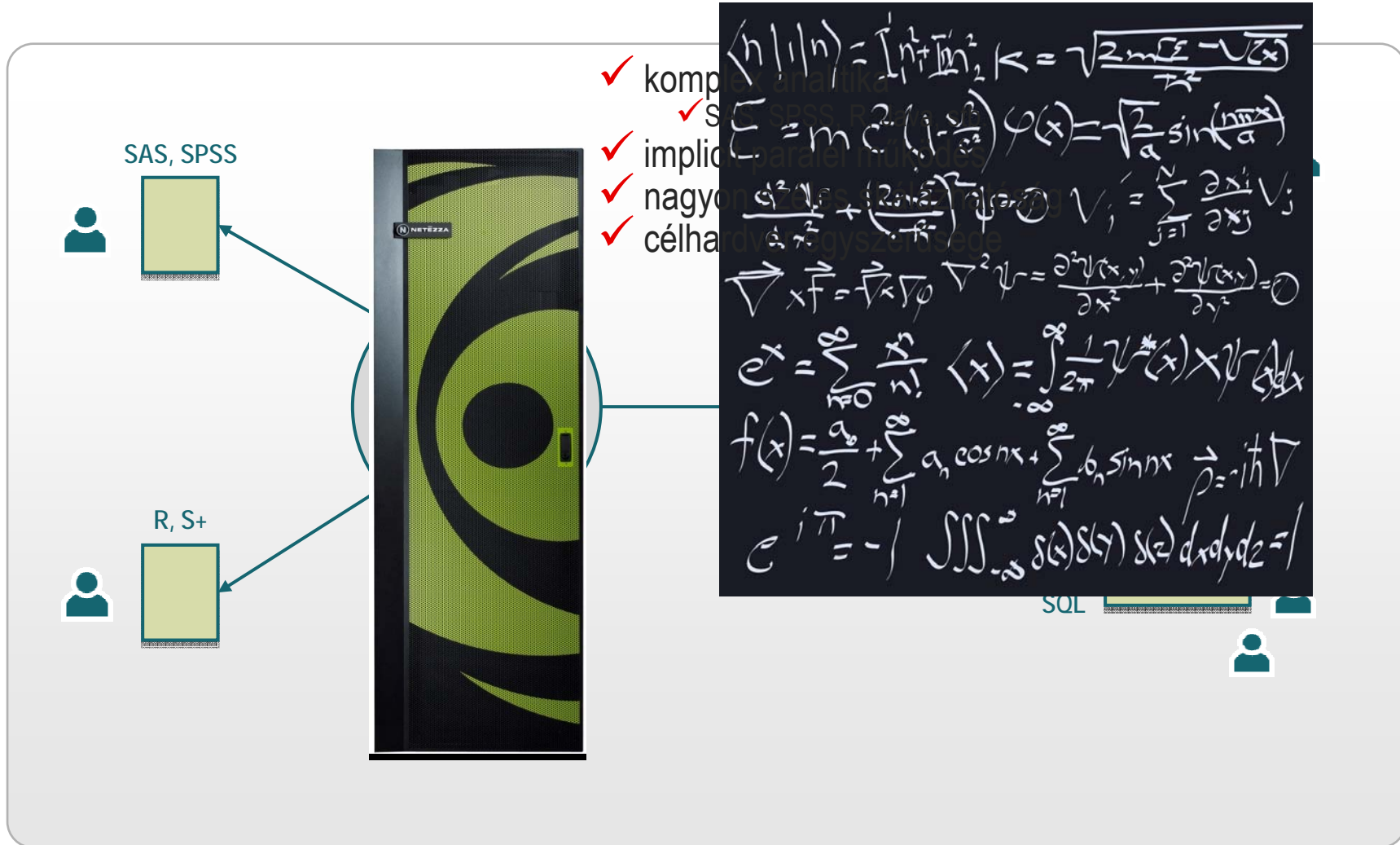
SQL ODBC JDBC OLE-DB



## Netezza integráció – LEKÉRDEZÉS



# Netezza analitika



## Analitikai függvények: nzStarter Kit

Data Mining	Predictive Analytics		Spatial
<p><b>Association Rules Mining</b></p> <p><b>Association</b></p> <ul style="list-style-type: none"> <li>• A Priori</li> <li>• FP-Growth</li> </ul> <p><b>Clustering</b></p> <p><b>K-Means</b></p> <p><b>Hierarchical Clustering</b></p> <ul style="list-style-type: none"> <li>• Divisive Clustering</li> <li>• Agglomerative Clustering</li> </ul> <p><b>Feature Extraction</b></p> <p><b>Dimension Reduction</b></p> <ul style="list-style-type: none"> <li>• Principal Components Analysis</li> </ul>	<p><b>Sample Size</b></p> <p><b>One-Way ANOVA</b></p> <ul style="list-style-type: none"> <li>• Complete Randomized Design</li> <li>• Randomized Block Design</li> </ul> <p><b>Regression</b></p> <p><b>Linear Regression</b></p> <p><b>Classification</b></p> <p><b>Decision Trees</b></p> <ul style="list-style-type: none"> <li>• Entropy Decision Tree</li> <li>• Gini Index Decision Tree</li> <li>• Regression Tree</li> </ul> <p><b>Neighborhood Methods</b></p> <ul style="list-style-type: none"> <li>• K Nearest Neighbors</li> </ul>	<p><b>Bayesian Methods</b></p> <p><b>Classifier</b></p> <ul style="list-style-type: none"> <li>• Naive Bayes</li> </ul> <p><b>Graphical Model</b></p> <ul style="list-style-type: none"> <li>• Bayesian Networks</li> </ul> <p><b>Model Testing</b></p> <p><b>Error Calculation</b></p> <ul style="list-style-type: none"> <li>• Cross Validation</li> <li>• Percentage Split</li> <li>• Train / Test</li> </ul>	<p><b>Geometric Functions</b></p> <p><b>Geometric Information</b></p> <p><b>Geometric Object Manipulation</b></p> <p><b>Geometric Analytics</b></p> <p><b>Conversion</b></p> <p><b>Comparison</b></p> <p><b>Distance and Area</b></p>



## Netezza megoldások



- **Adattárház:**
  - Skimmer; Twin Fin
  - Spatial option
  - Cloud
- **Analitika:** In database analytics :I -class
- **Kulcsrakész megoldás:**
  - Netezza Retail Analytic Appliance
- **Netezza Migrator:**
  - Adattárház migráció Oracle -> Netezza

# Netezza skálázhatóság



Skimmer	TwinFin	Cruiser
Dev & Test System	Data Warehouse High Performance Analytics	Queryable Archiving Back-up / DR
1 TB to 10 TB	1 TB to 1.5 PB	100 TB to 10 PB



## Bizonyítunk: megvalósíthatósági tanulmány (On-Site Proof of Concept)

- Bemutatjuk mennyire könnyed
- Bemutatjuk a teljesítményt
- Ügyfél saját környezetében mutatjuk be
- Integráció a meglévő rendszerekkel (ETL)
- Megmutatjuk az analitika gyorsítást
- Kiszámítjuk mennyi a várható TCO
- Megmutatjuk az üzleti értéket



## Napirend

- Adattárház helye az informatikai rendszerben
  - IBM Warehouse megoldások: Teljeskörű integráltság
  - „Egyszerűségtől az egyedi megoldásokig”
  - Tradicionális adattárház
  
- Optimalizált adattárház rendszerek
  - Netezza adattárház célszámítógép / célhardver
    - Mikor jó a Netezza
    - „Célhardver egyszerűsége”
    - IBM Netezza TwinFin architektúra
    - Netezza skálázhatóság
  
  - **Flexibilis adattárház és analitika: Smart Analytics System**
    - **Hatékony adattárház**
    - **Moduláris felépítés**
    - **Nagy rendelkezésre állás, nagyvállalti megoldások**
  
- Kérdések, válaszok

## Egyszerűség, rugalmasság, a választás szabadsága

### IBM Data Warehouse & Analytics Solutions

IBM Netezza	IBM Smart Analytics System	IBM InfoSphere Warehouse
Célhardver	Integrált rugalmas	Egyedi megoldás
<ul style="list-style-type: none"> <li>• Speciális adattárházra, ad-hoc analízisre optimalizált, 10x – 100X sebesség</li> <li>• Előtérben az analitikai lekérdezések</li> <li>• Az eszközök nem testreszabhatók, fizikai skálázhatóság</li> <li>• IBM, előkészített megoldás</li> <li>• Minimális adminisztráció, DBA, ügyfélerőforrás</li> </ul>	<ul style="list-style-type: none"> <li>• IBM által összehangolt szoftver és hardver csomag (fejlett DB2, Cognos funkcionalitás)</li> <li>• Komponenspaletta</li> <li>• Rugalmas megoldás különböző adat, felhasználói modulok kezelésére (egyaránt használható OLTP-re, DW-re)</li> <li>• Fizikai és logikai skálázhatóság (modulok, hangolás)</li> <li>• Ár/teljesítmény „jóméretesíthető”</li> </ul>	<ul style="list-style-type: none"> <li>• Ügyféllel közösen kialakított egyedi megoldás</li> <li>• Szerverek, tároló eszközök, adatbázisok és analitikai szoftverek széles választéka</li> <li>• Mind IBM, mind nem IBM megoldások egyben</li> </ul>



## Flexibilis adattárház és analitika

- **Analitika**
  - Üzleti Intelligencia (Cognos)
  - Adatkocka alapú analitika
  - Adatbányászat, text analitika
- **Hatékony adattárház**
  - Adattárházra optimalizált platform
  - Fejlett terhelésmenedzsment
  - Automatizáció
- **Hardver és szolgáltatások**
  - Rugalmas szerverplatform
  - Moduláris felépítés
  - Kiépítés, telepítés, support, prémium szolgáltatások

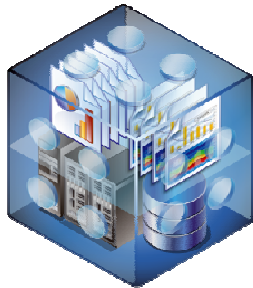


Nyers adatokból → üzleti érték

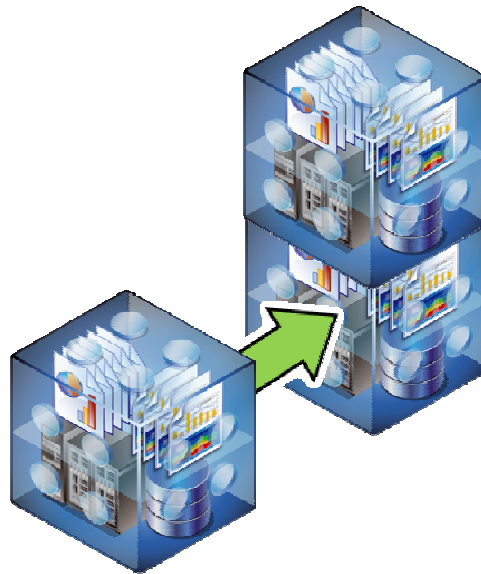


## Nagyon széles határok közötti bővíthetőség

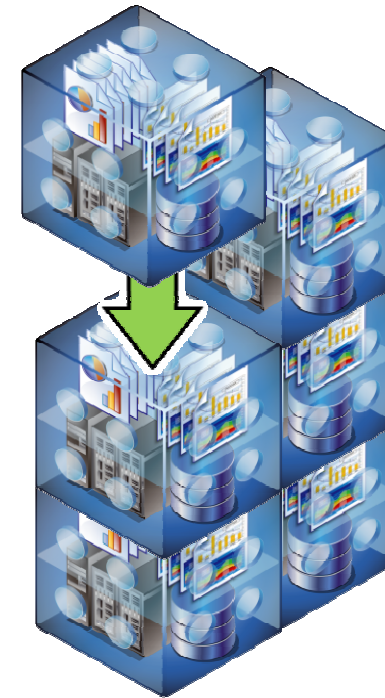
Alaprendszer



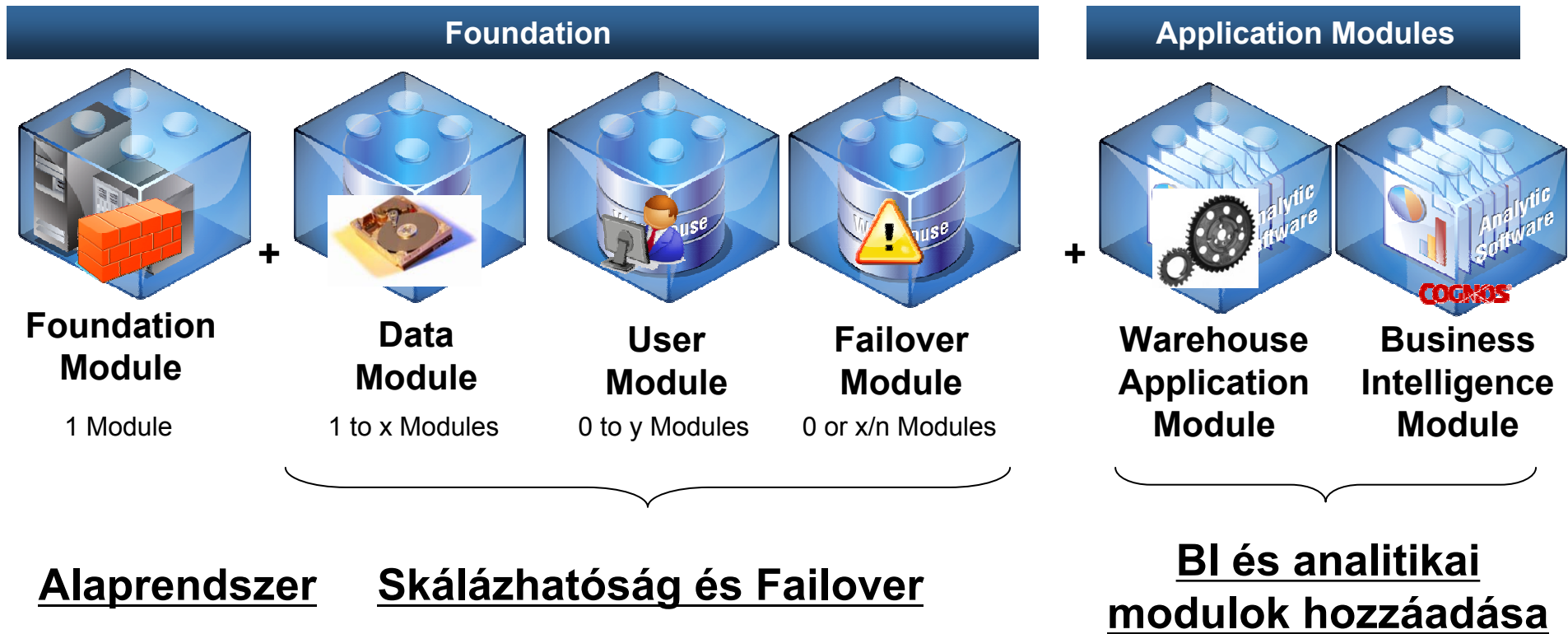
Kapacitás bővítés



Funkcionális bővítés



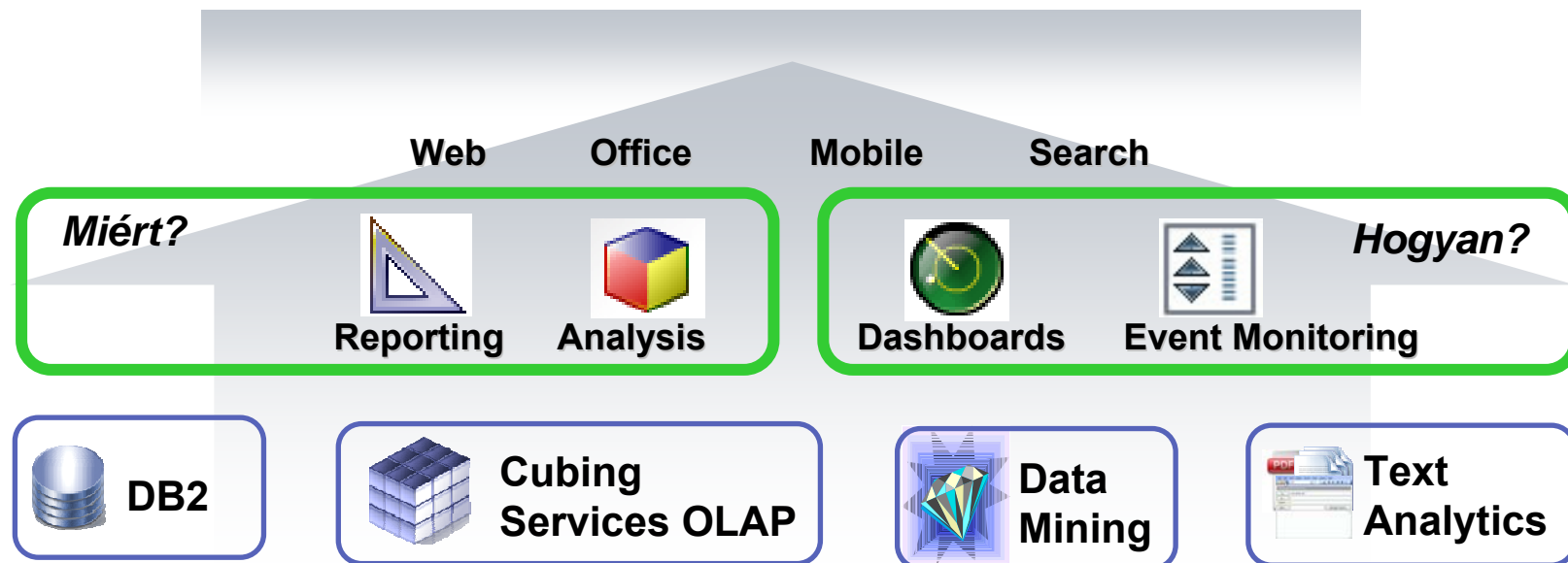
## IBM Smart Analytics System: a moduláris architektúra



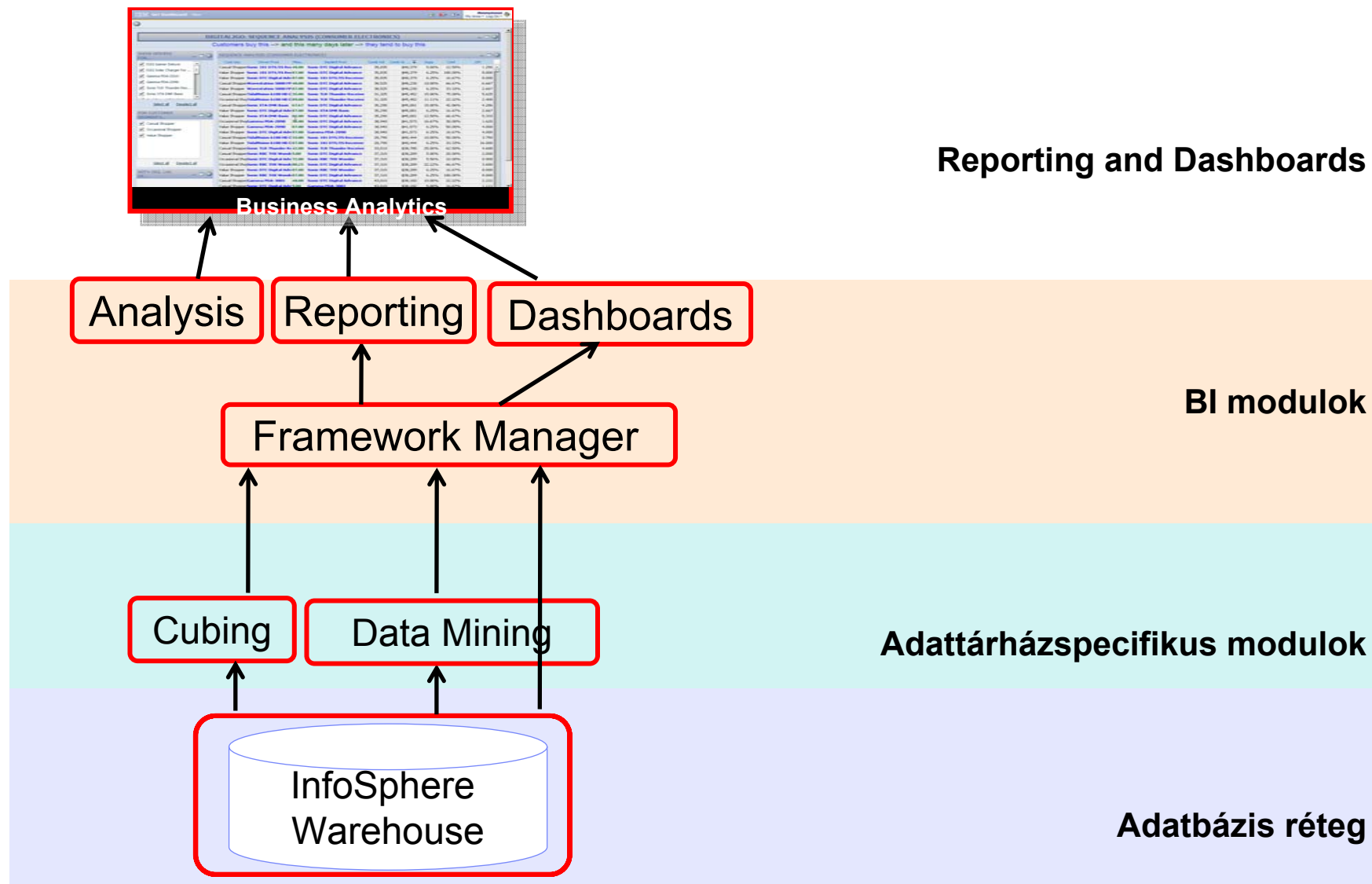
Rendszer fejlesztés moduláris útjai →

## Hatékony vállalati analitika

- BI eszközök széles tára
- Növekvő ügyfél igényekhez igazodó
- Nagy rendelkezésre állású
- Integrált OLAP, adatbányászati, természetes nyelvi analitikai eszközkészlet



# IBM Smart Analytics System: integrált megoldáshalmaz



## Termékvonal termékcsalád



### Compact Systems

- Kompakt rendszer
- Nem moduláris HW architektúra

Mid-Market, Departments and Sandbox /  
TestDev



### Modular Systems

- Tetszőleges adatmennyiség
- Lineáris skálázhatóság

Nagyvállalati rendszerek

## Kérdések, válaszok



KÖSZÖNÖM

A FIGYELMET!

[lpakozdi@hu.ibm.com](mailto:lpakozdi@hu.ibm.com)

**IBM**

**IBM**  
**Co.**

**BUSINESS**  
**INTERNATIONAL**  
**MACHINES**

**IBM**