IBM Information Management

IBM Dynamic Warehouse

Николай Куликов
Nikolay_Kulikov@ru.ibm.com

March 2007

© 2007 IBM Corporation

Ясная, необходимая информация в Любом месте в Любое время

Information On Demand
- Оптимизация каждой транзакции
- Call Центры, Сотрудники в полях

Помощь в раскрытии преступлений - предоставление необходимой информации в момент прибытия офицеров на место происшествия

OLAP & Data Mining
- Запасы, Товары, Движение

Оптимизация работы полицейских

Запросы и Отчеты
- Финансы, Продажи

Статистика преступлений

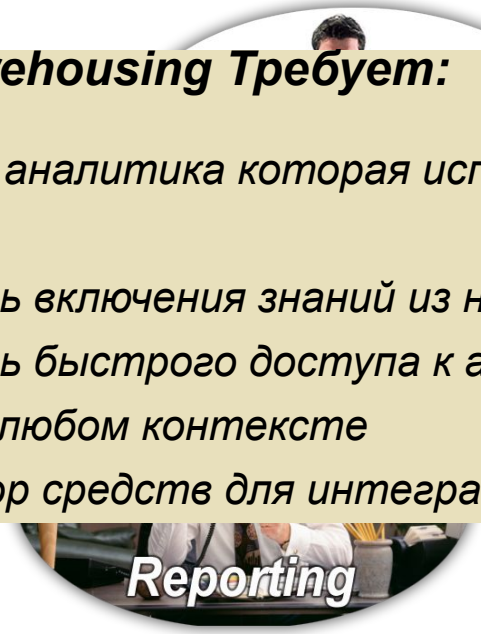
© 2007 IBM Corporation

Dynamic Warehousing

A New Approach to Leveraging Information

Information On Demand
Оптимизация реальных процессов

Dynamic Warehousing



ОЦ

реко

Зап
Дл
Что произошло

Dynamic Warehousing Требует:

1. Встроенная аналитика которая используется как часть бизнес процесса
2. Возможность включения знаний из неструктурированной информации
3. Возможность быстрого доступа к агрегированной и очищенной информации в любом контексте
4. Полный набор средств для интеграции информации

© 2007 IBM Corporation

Больше примеров Dynamic Warehousing в действии

Традиционные Хранилища	Dynamic warehousing
Аналитика и отчетность мошенничеству в страховании	Определение потенциально мошеннических заявок до согласования и оплаты ***Преобразование страхования***
Отчетность по претензиям клиентов	Выявление потенциальных похожих заявок, снижение ухода клиента, возможность кросс продаж ***Преобразование клиентской службы***
Анализ и отчетность по продажам за прошедший период	Выявление необходимой информации о клиенте для идентификации возможности дополнительных продаж ***Преобразование эффективности продаж***
Статистика и отчетность по преступлениям	Выявление похожих инцидентов до прибытия на место происшествия ***Преобразование борьбы с преступниками***

© 2007 IBM Corporation

Вызовы стоящие перед традиционными ХД

ХД должны сейчас:

- **Адресованы** к расширяющимся потребностям аналитики по запросу
- **Использовать** ВСЕ типы информации включая неструктурированную
- **Обслуживать** увеличивающиеся количество и типы приложений и пользователей с различными потребностями уровня обслуживания

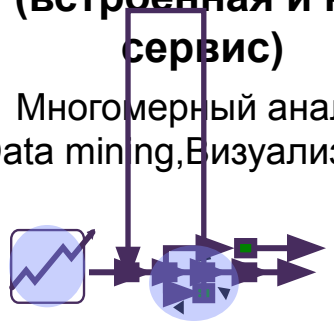
Увеличивающаяся Смешанная нагрузка и
Постоянно изменяющиеся требования различных бизнес направлений
требует более Динамичных Возможностей ХД

© 2007 IBM Corporation

IBM предоставляет больше чем хранилище

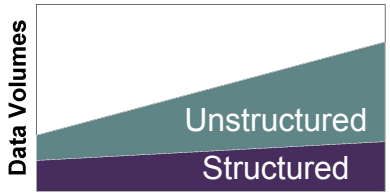
Встроенная Аналитика (встроенная и как сервис)

Многомерный анализ
Data mining,Визуализация



Традиционное ХД

За границами традиционных структурированных данных



"As a direct effect of the mixed workload, with continuous loading and the increase in automated transactions from the functional analytics in OLTP, the transactional DBMSs have an edge that challenges the DW DBMSs (such as Teradata)"

Gartner Data Warehouse Magic Quadrant, 2006

Преим
транзак
осно

Оптимизация для быстрого доступаысокая доступность и надежность,

Масштабируемость, безопасность, аудируемость

диски, лучшая утилизация дисков, повышение производительности запросов

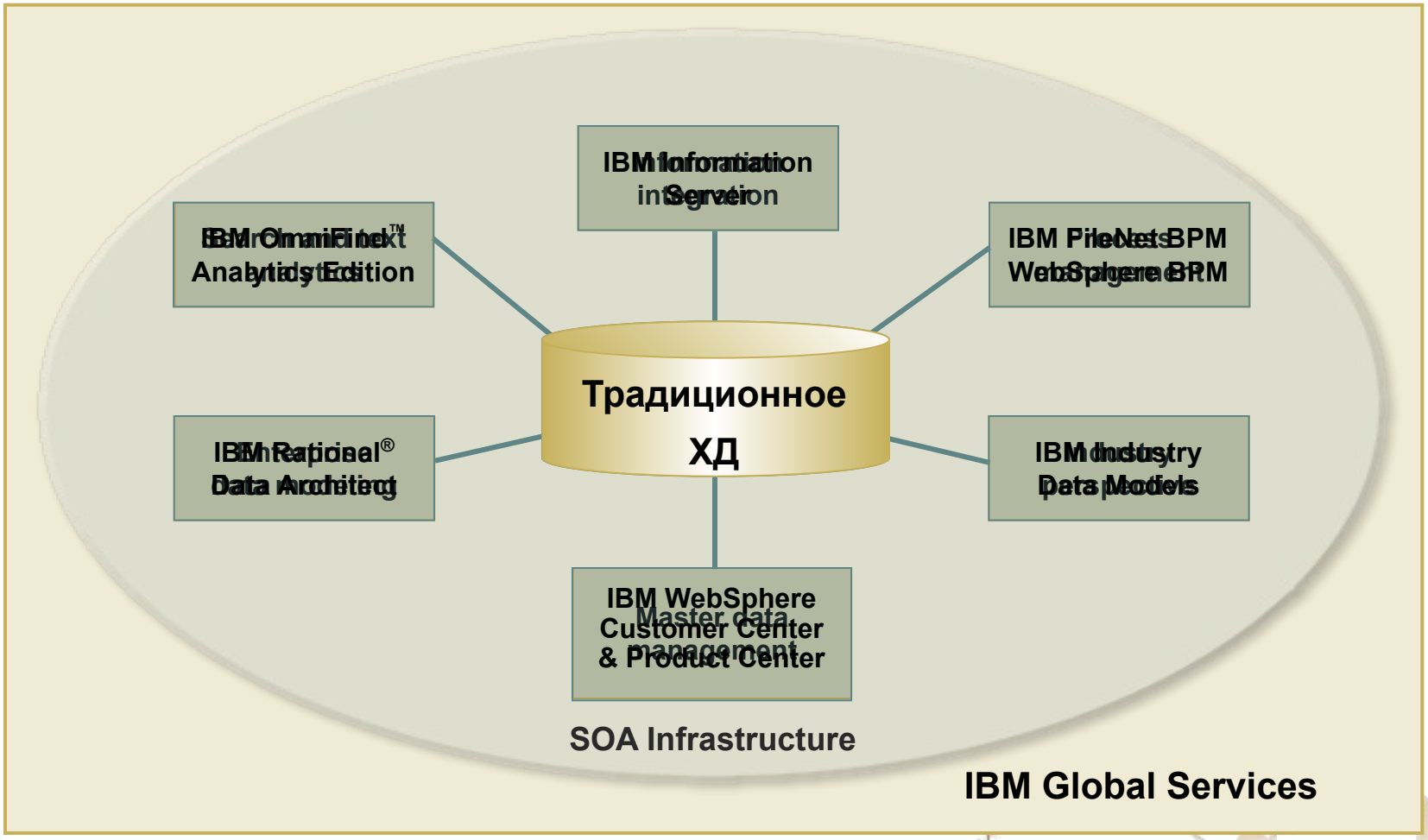
ные

ура
Shared-nothing
Партиционирование
Управление нагрузкой

© 2007 IBM Corporation

IBM Dynamic Warehousing

Интегрированное предложение



© 2007 IBM Corporation

Решение от IBM: IBM Information Server

Предоставление информации которой вы можете доверять

IBM Information Server

Общее Внедрение

Понимание	Очистка	Трансформация	Доставка
Обнаружение, моделирование, управление информацией	Стандартизация, объединение, и корректировка информации	Объединение и реструктуризация информации для нового пользования	Синхронизация, виртуализация, перемещение информации

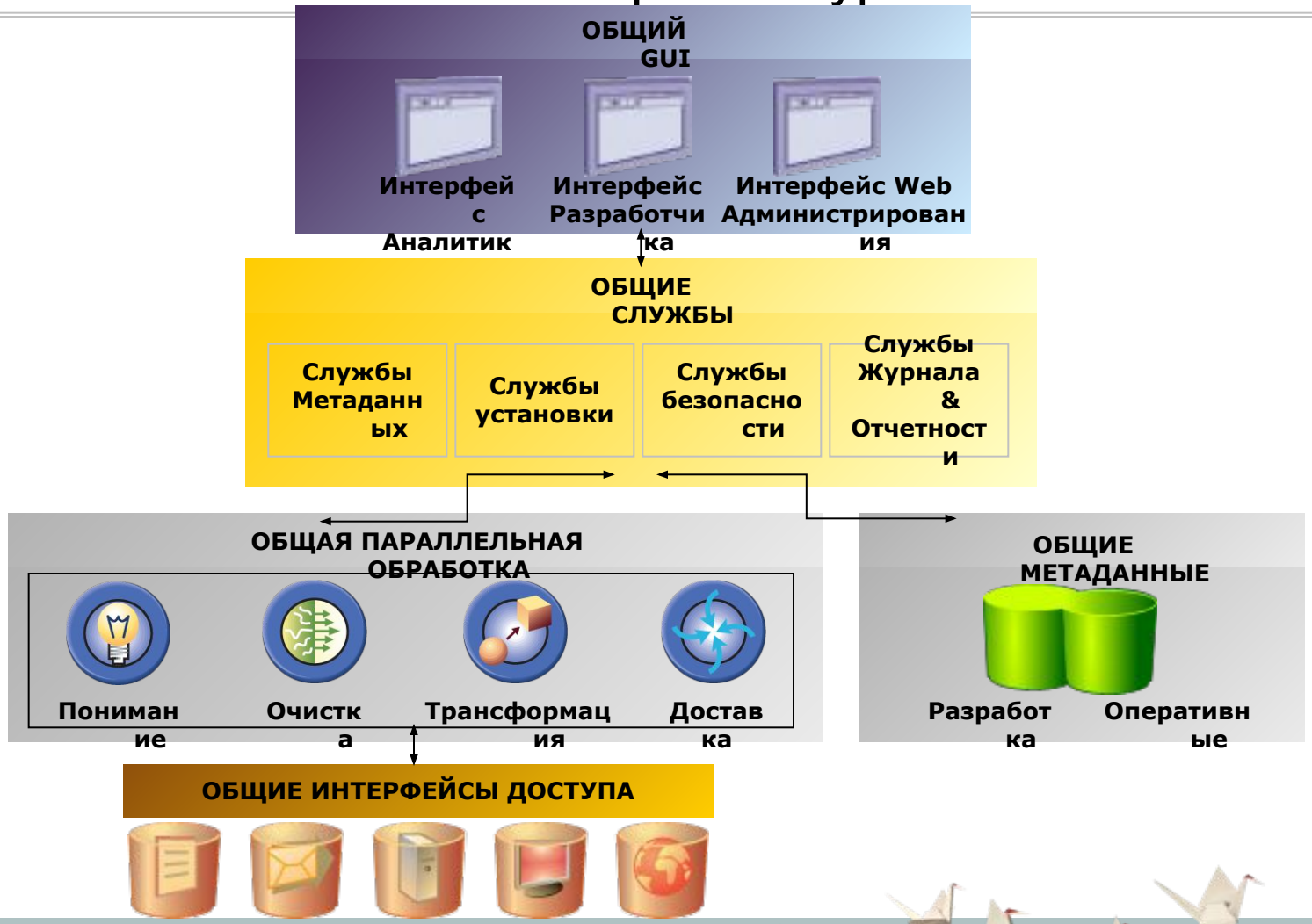
Общее управление метаданными

Параллельная обработка

Широкие возможности доступа к Приложениям, Данным, и Контенту

© 2007 IBM Corporation

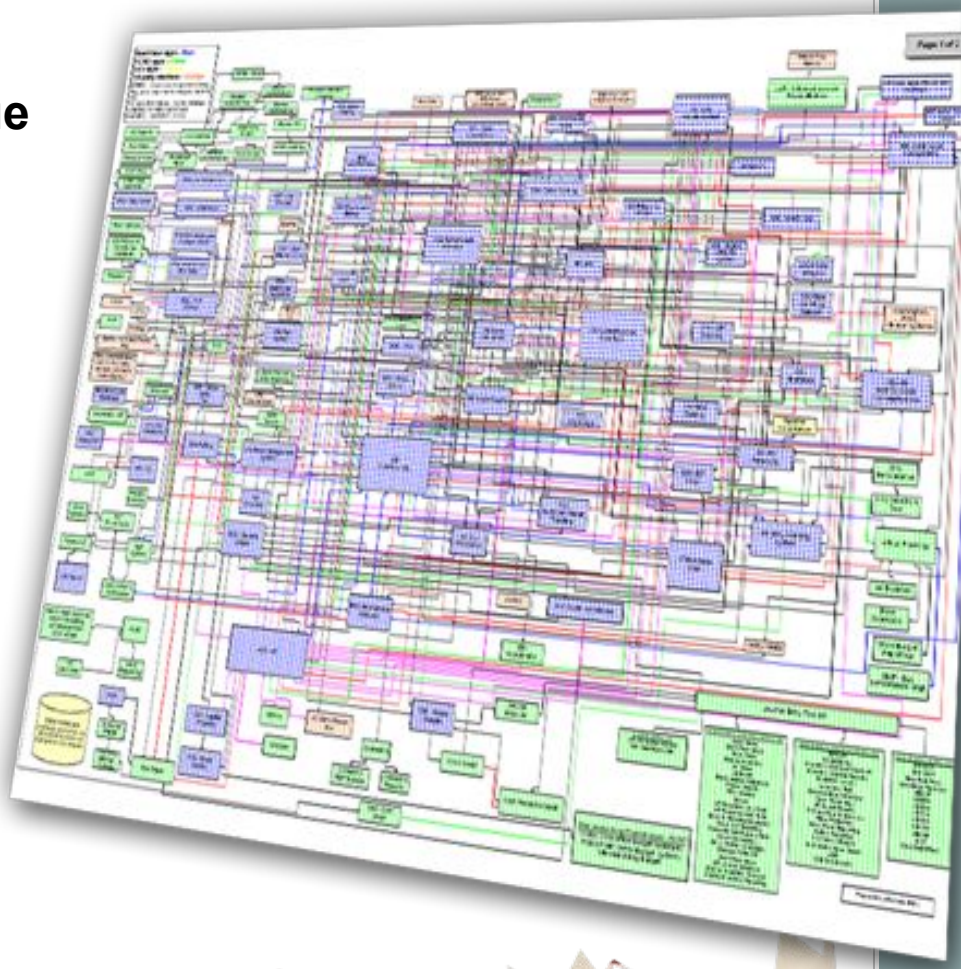
IBM Information Server - Архитектура



© 2007 IBM Corporation

Почему важно начинать с Понимания?

- **Где необходимая информация?**
- **Как я могу ее получить если она мне нужна?**
- **Что она означает?**
- **Могу ли я ей доверять?**
- **Когда я ее получаю в какой форме она мне нужна?**
- **Как я могу ее контролировать?**

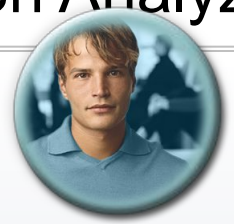


© 2007 IBM Corporation

Физические метаданные: IBM WebSphere Information Analyzer

- Основанный на данных анализ приложений, БД и файловых источников данных
- Безопастная, детальная профилировка полей, между полями и между источниками
- Создание метаданных как результатов профилирования
- Результаты постоянно доставляются посредством IBM Information Server

Subject Matter Experts

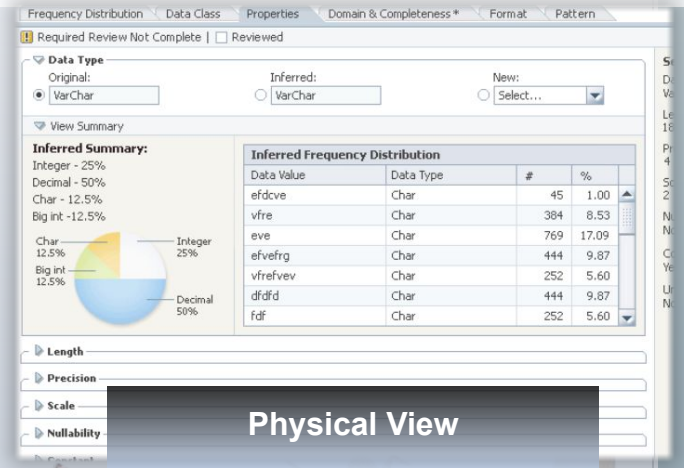


Data Analysts

Understand

IBM WebSphere Information Analyzer

Анализ структур исходных данных и мониторинг их качества и соответствия правилам



© 2007 IBM Corporation

Бизнес метаданные: IBM WebSphere Business Glossary

- Построенная на Web система для описания, управления & общего доступа к бизнес метаданным
- Выравнивание действий IT с целями бизнеса
- Предоставление бизнес контекста к информационным активам

Subject Matter Experts

Business Users

Understand

IBM WebSphere Business Glossary

Create and manage business vocabulary and relationships, while linking to physical sources

Business View

Database = DB2

Schema = NAACCT

Table = DLYTRANS

Column = ACCT_NO

data type = char(11)

Technical

Business

GL Account Number

The ten digit account number. Sometimes referred to as the account ID. This value is of the form L-FIIIIVVVV.

© 2007 IBM Corporation

Очистка данных: IBM WebSphere QualityStage

- Специализированные функции очистки данных прозрачно интегрированы с DataStage
- Визуальные инструменты для определения комплексных отображений и логики сохранения полей
- Проверка очистки, стандартизации, де-дубликации информации
- Единая версия правды

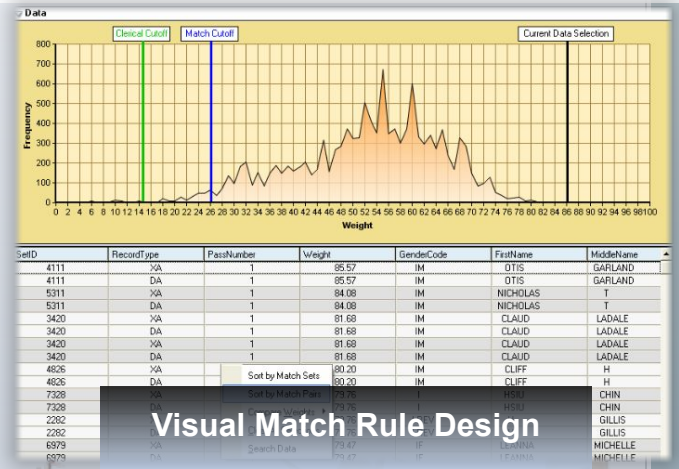
Subject Matter Experts

Data Analysts

Cleanse

IBM WebSphere QualityStage™

Standardize and correct source data fields, and match records together across sources to create a single view

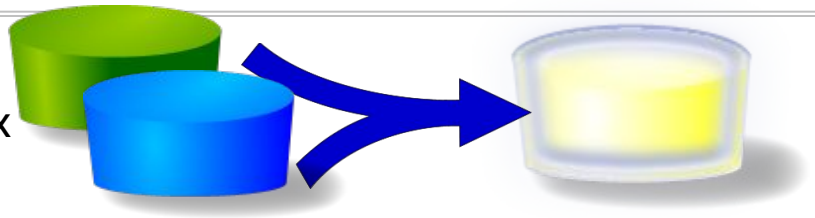


Visual Match Rule Design

© 2007 IBM Corporation

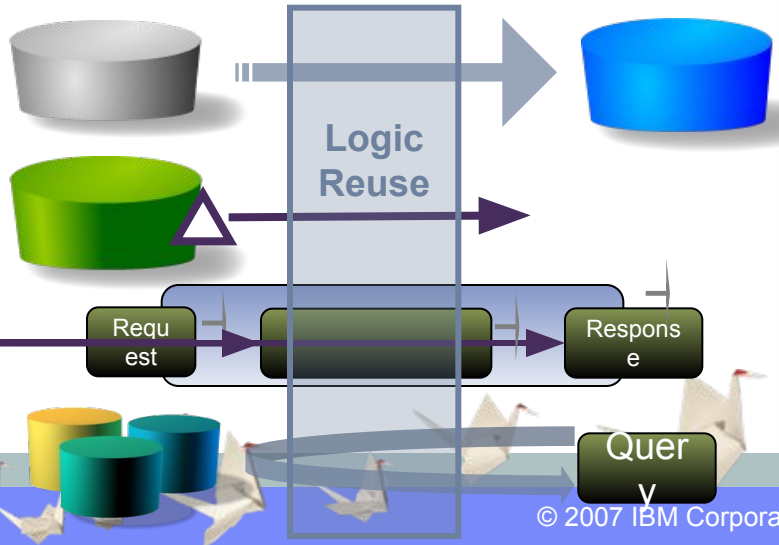
Что важно в трансформации и доставке данных?

- Трансформация это ключ для предоставления информации для использования ее в новых бизнес контекстах – это требует что бы эта информация была основанная на метаданных



- Разработана для использования экспертами по информации и управляется посредством метаданных

- Трансформация и доставка может быть использована множеством механизмов
  - Перемещение больших объемов заданиями (batch)
  - Ответ на событие в реальном времени
  - SOA
  - Федеративные запросы



© 2007 IBM Corporation

Трансформация & Перемещение: IBM WebSphere DataStage

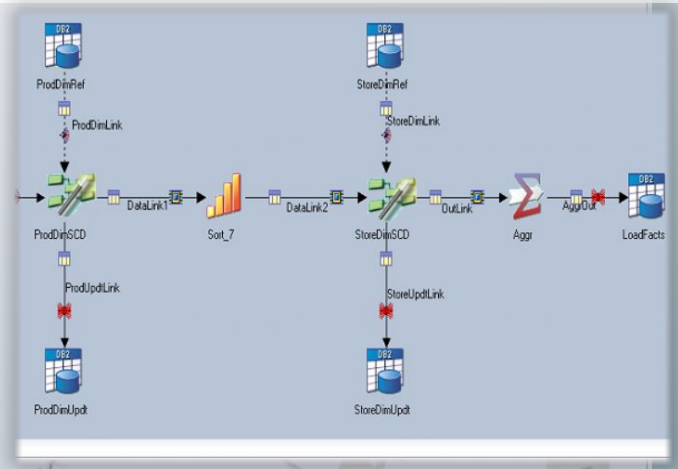
- Визуальная разработка потоков данных с сотнями встроенных трансформаций
- Оптимизированное использование объектов интеграционных
- Возможность параллельной обработки без необходимости изменять процессы обработки
- Возможность работы как с batch так real-time операциями

Transform

Deliver

IBM WebSphere DataStage®

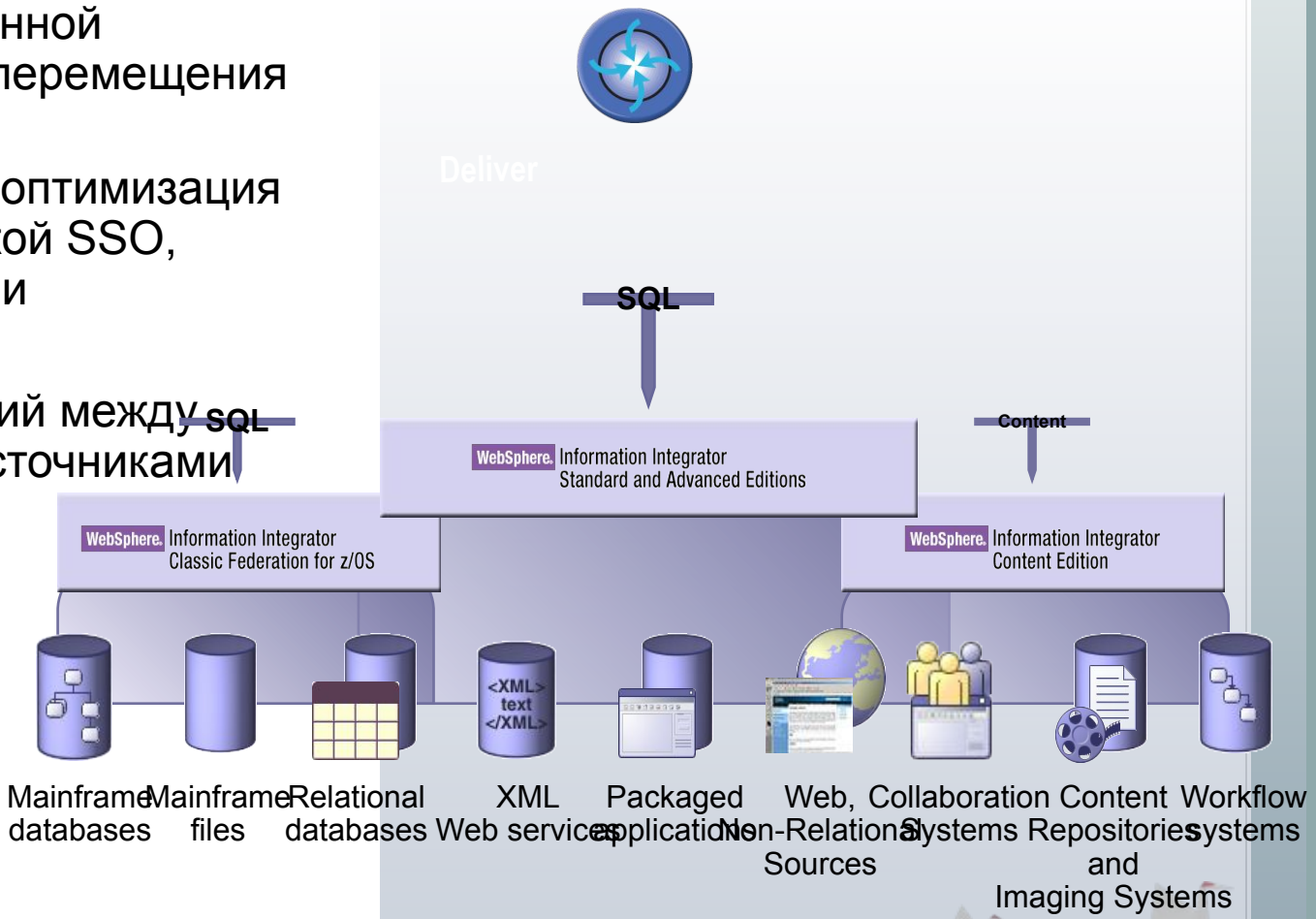
Трансформация и агрегация любых объемов информации



© 2007 IBM Corporation

Федеративные данные: IBM WebSphere Federation Server

- Доступ к распределенной информации без ее перемещения
- Лучшая в индустрии оптимизация запросов с поддержкой SSO, функции компенсации
- Поддержка транзакций между распределенными источниками



© 2007 IBM Corporation

Федеративные запросы делают интеграцию простой как SQL

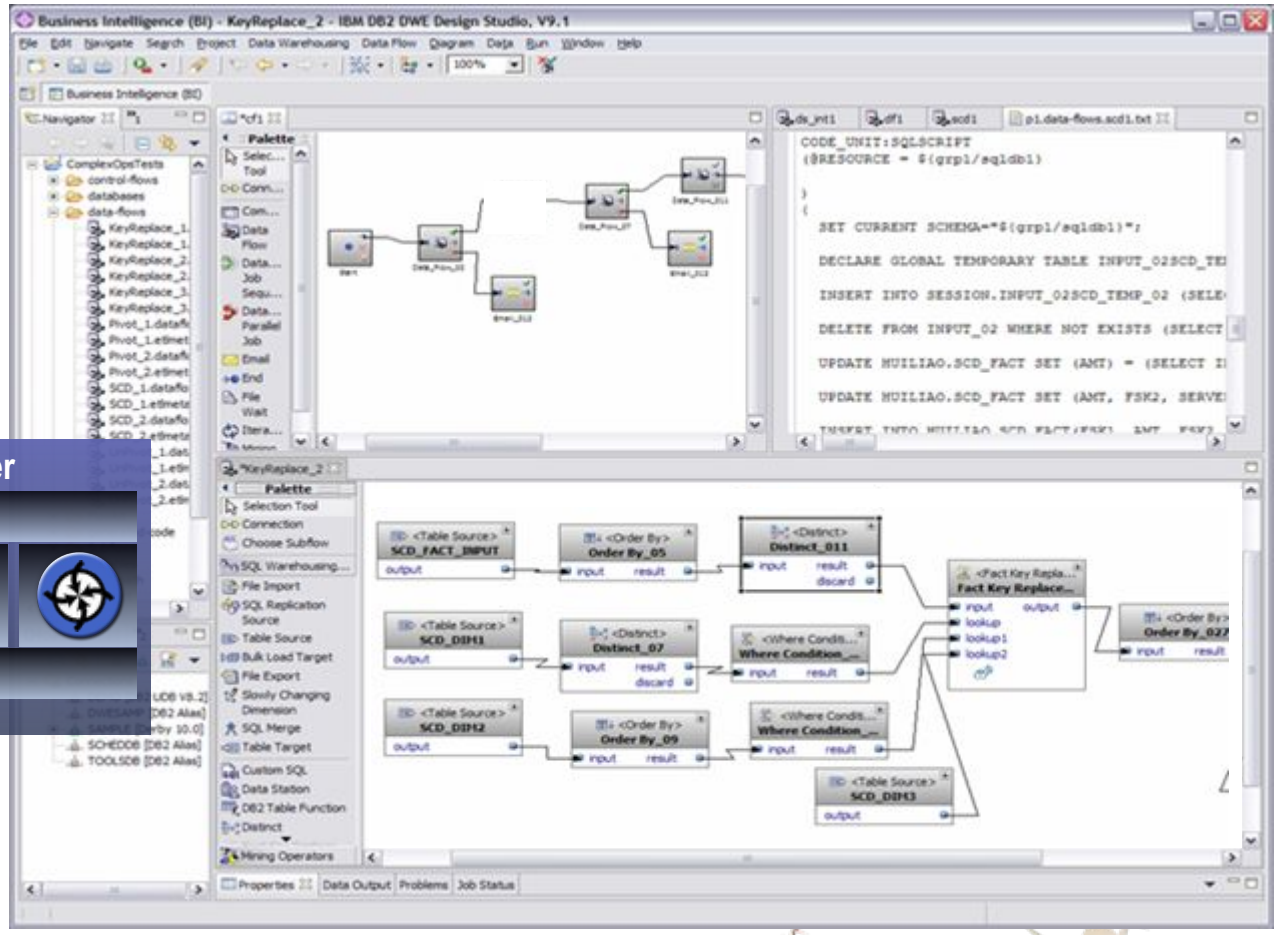
```
     parameters_return_billto_key as BILL_TO_KEY

parameters_return_shipto_key as SHIP_TO_KEY
CASES_SHIPPED
GROSS_SALES
URL

     GETKEYSSOAP_GETKEYSREALTIME_NN

GLOBAL_SALES_TRAN_NN




URL_INVOICES

     getkeysrealtime_ship_to_number
     getkeysrealtime_ship_to_number    URL_INVOICES.shipno
     ltrim(rtrim(translate(ship_to_number, ' ', x'0a')))
             getkeysrealtime_ship_to_number
     parameters_return_billto_key
     ltrim(rtrim(translate(sales_order_number, ' ', x'0a')))
             URL_INVOICES.orderno
```

Один запрос SQL :

- Web Service
- XML Documents
- Data Warehouse
- Unstructured Data

© 2007 IBM Corporation

Интегрированный набор инструментов для Dynamic Warehousing

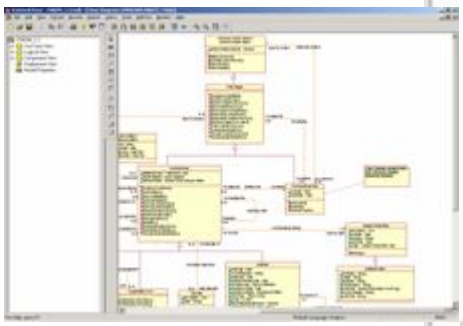


IBM Information Server

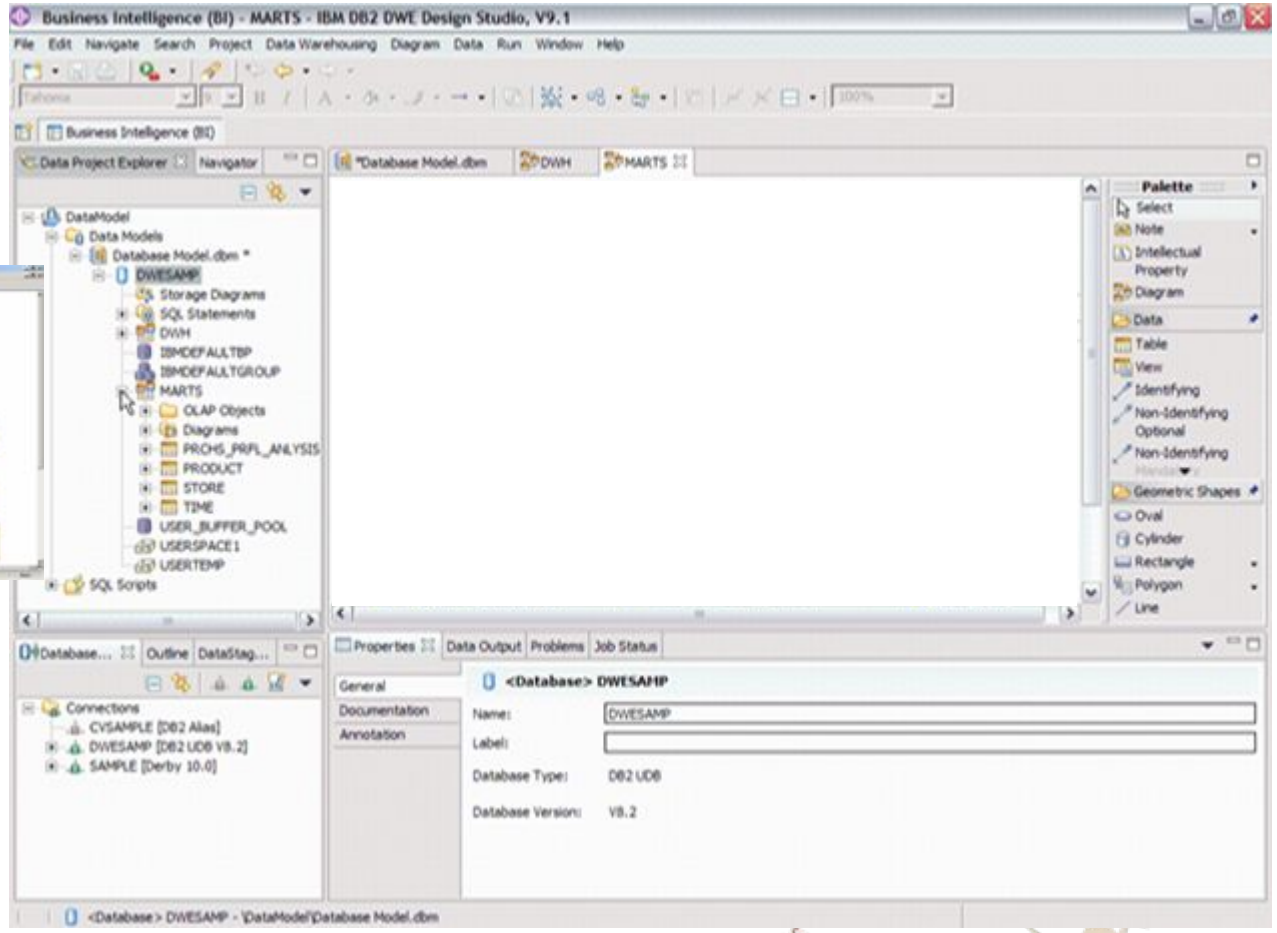
© 2007 IBM Corporation

Интегрированный набор инструментов для Dynamic Warehousing

Rational. software



Data Architect



© 2007 IBM Corporation

Introducing IBM Balanced Warehouse™

A fast track to warehousing

Простота

- Предопределенная конфигурация для упрощения комплекса
- Один номер для поддержки решения

Balanced Configuration Unit (BCU)

Преконфигурированный, Претестированный, масштабируемый набор Программного и Аппаратного Обеспечения

Гибкость при росте

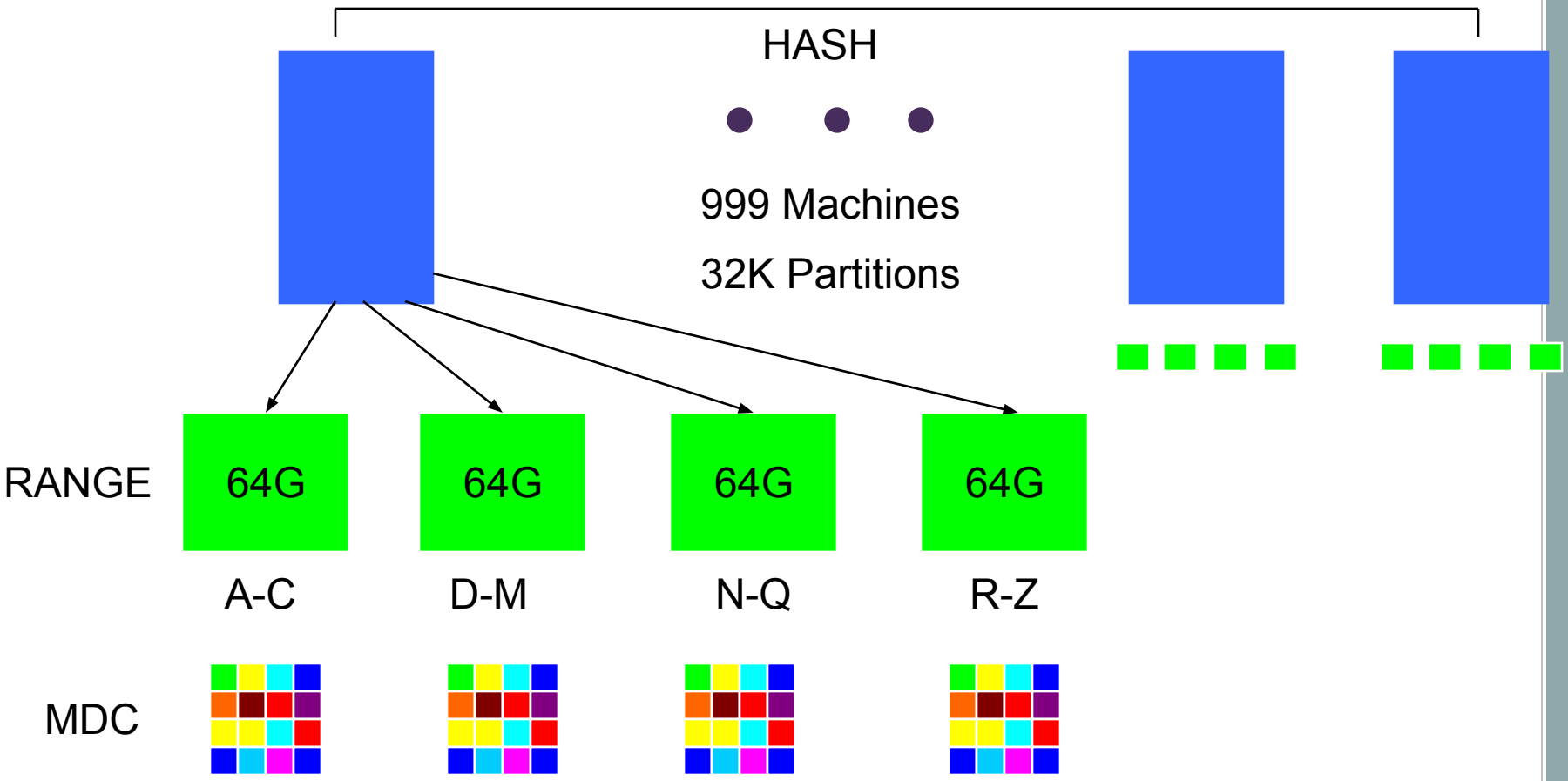
- Простое Добавление BCUs для увеличивающихся потребностей
- Надежно, возможность повторного использования для других задач

Оптимизирован для Производительности

- Преконфигурирован и сертифицирован для предоставления гарантированной производительности
- Основан на лучших практиках для уменьшения риска

© 2007 IBM Corporation

Гибридное партиционирование

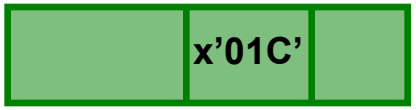


© 2007 IBM Corporation

Компрессия строк

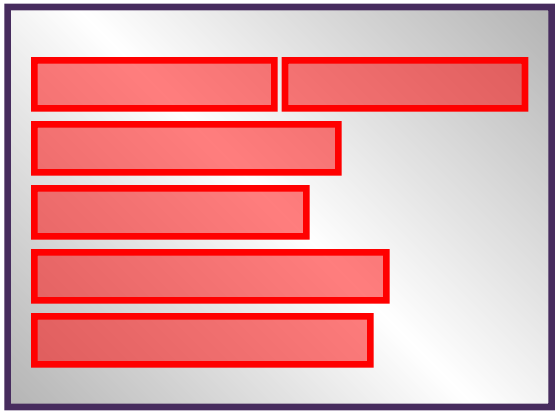
Без Компресии

С Компрессией

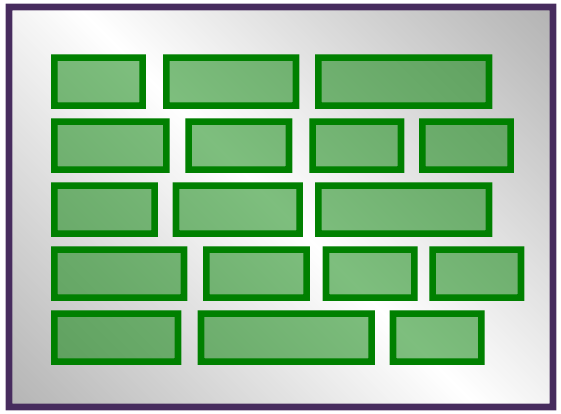


Общая последовательность заменяется на 12 байт

Страница без Компрессии



Страница с компрессией

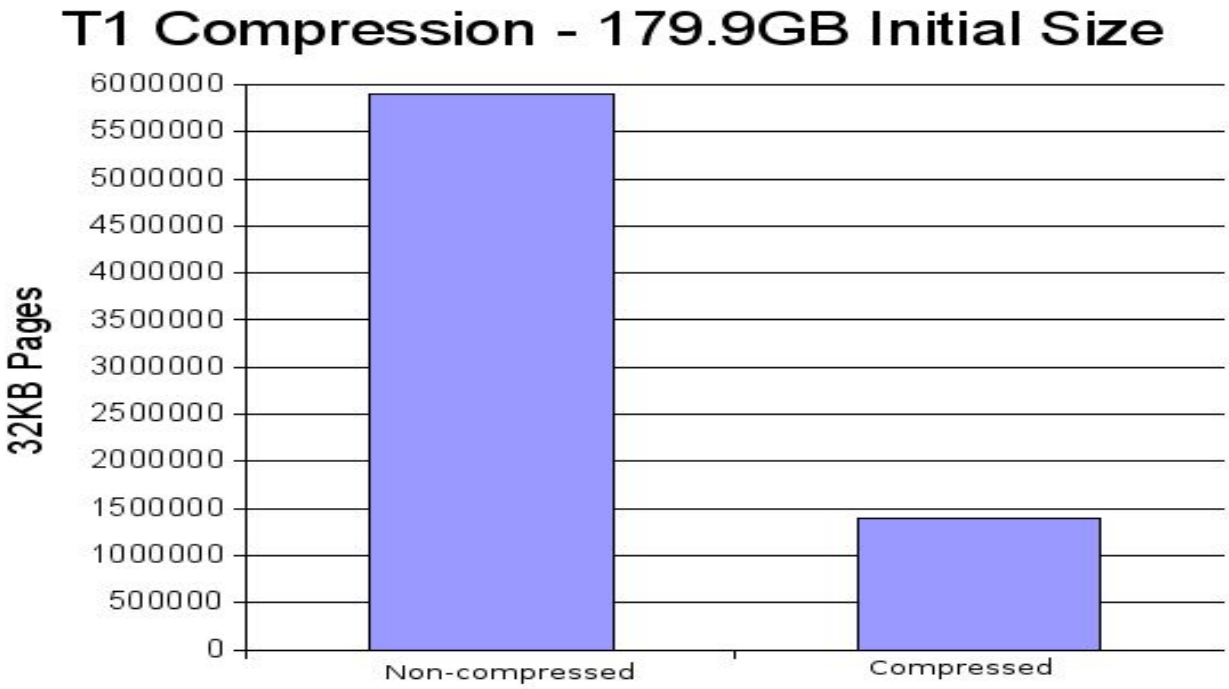


© 2007 IBM Corporation

DB2 - More Compression Ratios (Customer Data)

Compression Type	32KB Page Count	Space Required on Disk
No compression	5893888	179.9GB
Row compression	1392446	42.5GB

% Pages Saved: 76.4%



© 2007 IBM Corporation

Индустриальные Модели данных

BDW
(Banking Data Warehouse)

- Прибыльность
- Маркетинг
- Управление рисками
- Управление активами пассивами
- Соответствие Требованиям

Финансовые Рынки
(Financial Markets Data Warehouse)

- Управление Рисками
- Управление активами пассивами
- Соответствие требованиям

Управление Здоровьем
(Health Plan Data Warehouse)

- Заявки
- Medical management
- Управление сетью
- Продажи, маркетинг and membership
- Финансы

Страхование
(Insurance Information Warehouse)

- Клиенты
- Заявки
- Прибыльность
- Соответствия требованиям
- Управления рисками

Розница
(Retail Data Warehouse)

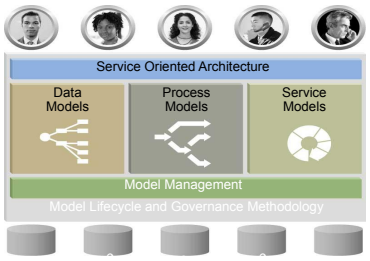
- Клиенты
- Управление закупками
- Оптимизация складов
- Управление цепочкой
- Соответствие требованиям

Телекоммуникации
(Telecommunications Data Warehouse)

- Управление оттоком
- Управление взаимоотношениями
- Продажи и маркетинг
- Качество обслуживания
- Профайл пользователя

© 2007 IBM Corporation

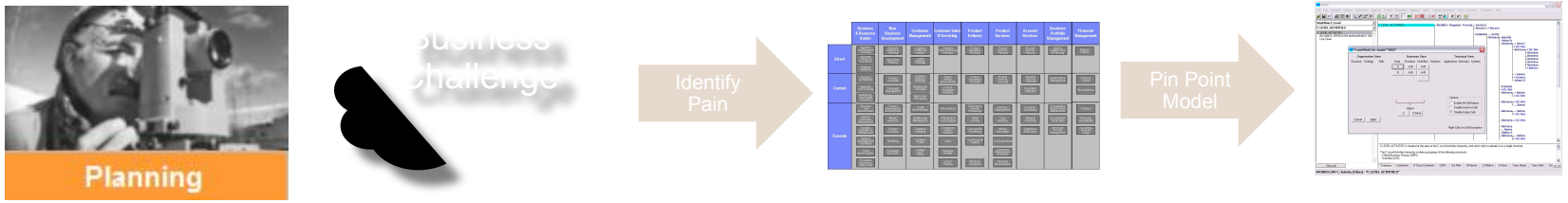
IBM Industry Data Models: Ключевые компоненты



- **Модель данных для терминов**
  - Словарь бизнес определений для единых определений в рамках организации и индустрии
- **Модель EDW**
  - Эта модель определяет как множество источников могут быть консолидированы в единую логическую структуру
- **Проектные представления**
  - Логическое подмножество EDW и BSTs для решения конкретной задачи
- **Шаблоны бизнес решений (BST)**
  - Набор Key Performance Indicators для создания OLAP кубов, витрин, шаблонов отчетности

© 2007 IBM Corporation

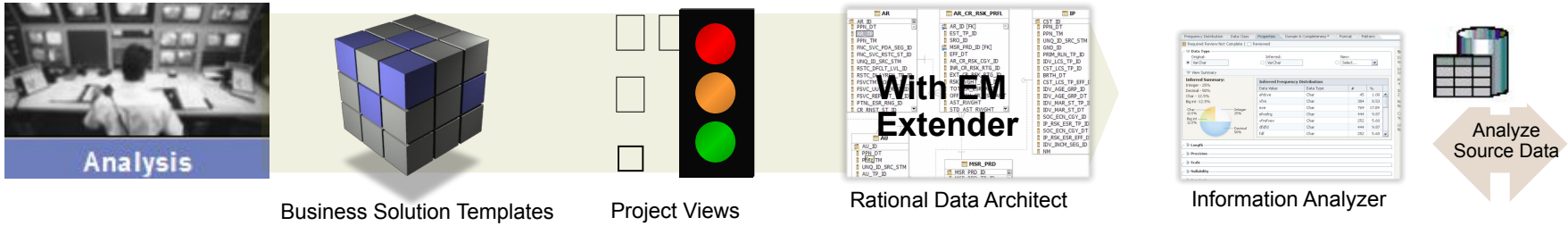
Фаза планирования



- Создание стандартов на данные и их управление
- Бизнес планирование и развития IT
- Стабильность проектов

© 2007 IBM Corporation

Фаза анализа



- RDA (Enterprise Model Extender)
  - Выявление Бизнес требований
  - Структурированная по области проблем
  - Определяет область для реализации IT реализации
- Дополнение Information Analyzer
  - Предоставление необходимого анализа данных

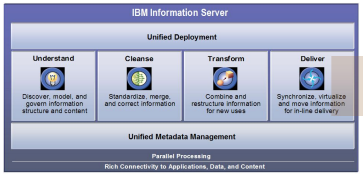
© 2007 IBM Corporation

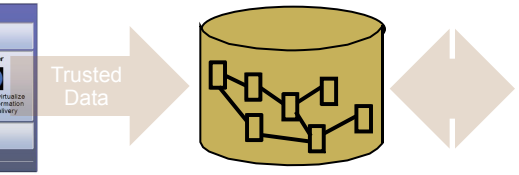
Создание дизайна

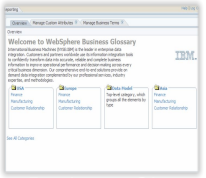
- RDA (plus Enterprise Model Extender)
  - Генерация ХД, витрин/кубов
  - Инкрементальная, проектная, корпоративная разработка
  - Трассировка бизнес требований до уровня кода
  - Сравнение, объединение моделей

© 2007 IBM Corporation

Фаза реализации



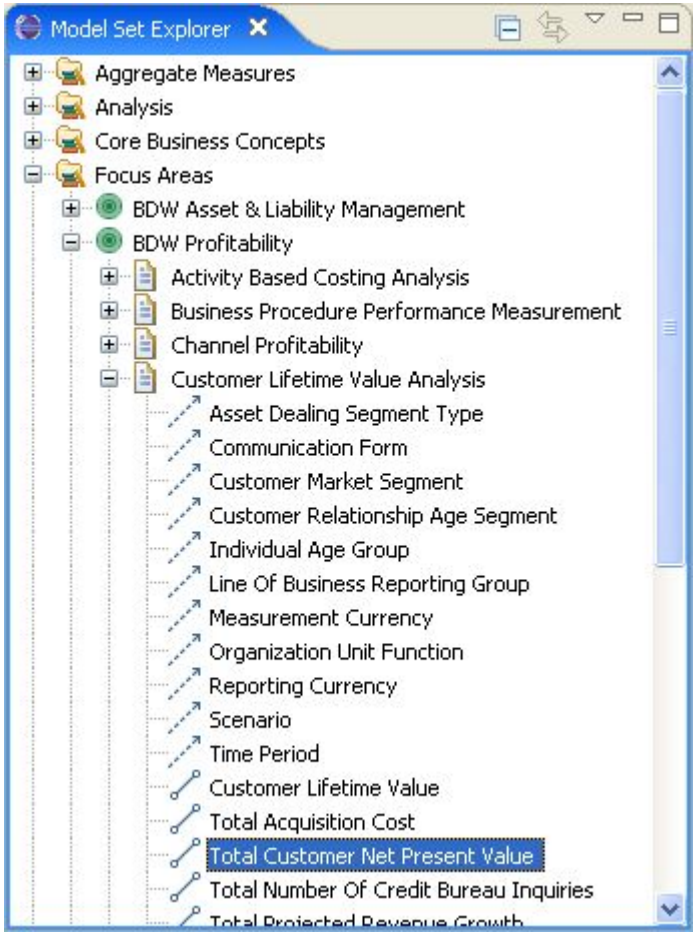


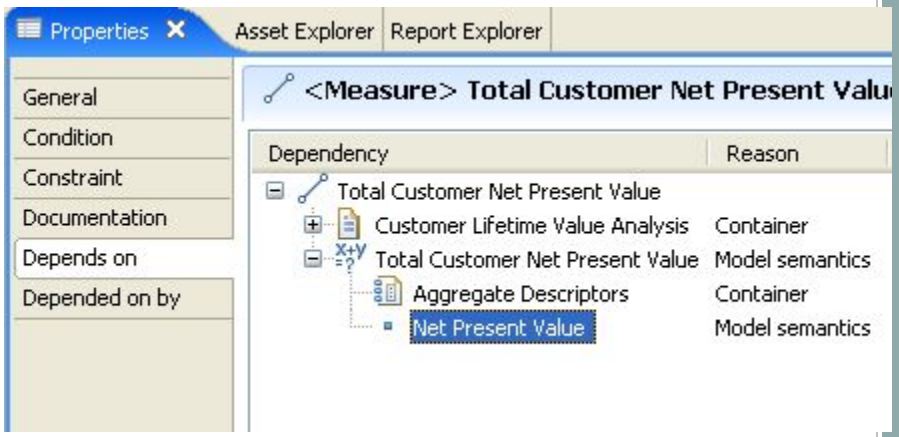


- Определение целевых структур
- Общий глоссарий
- Инкрементальная разработка без переписывания

© 2007 IBM Corporation

Анализ требований



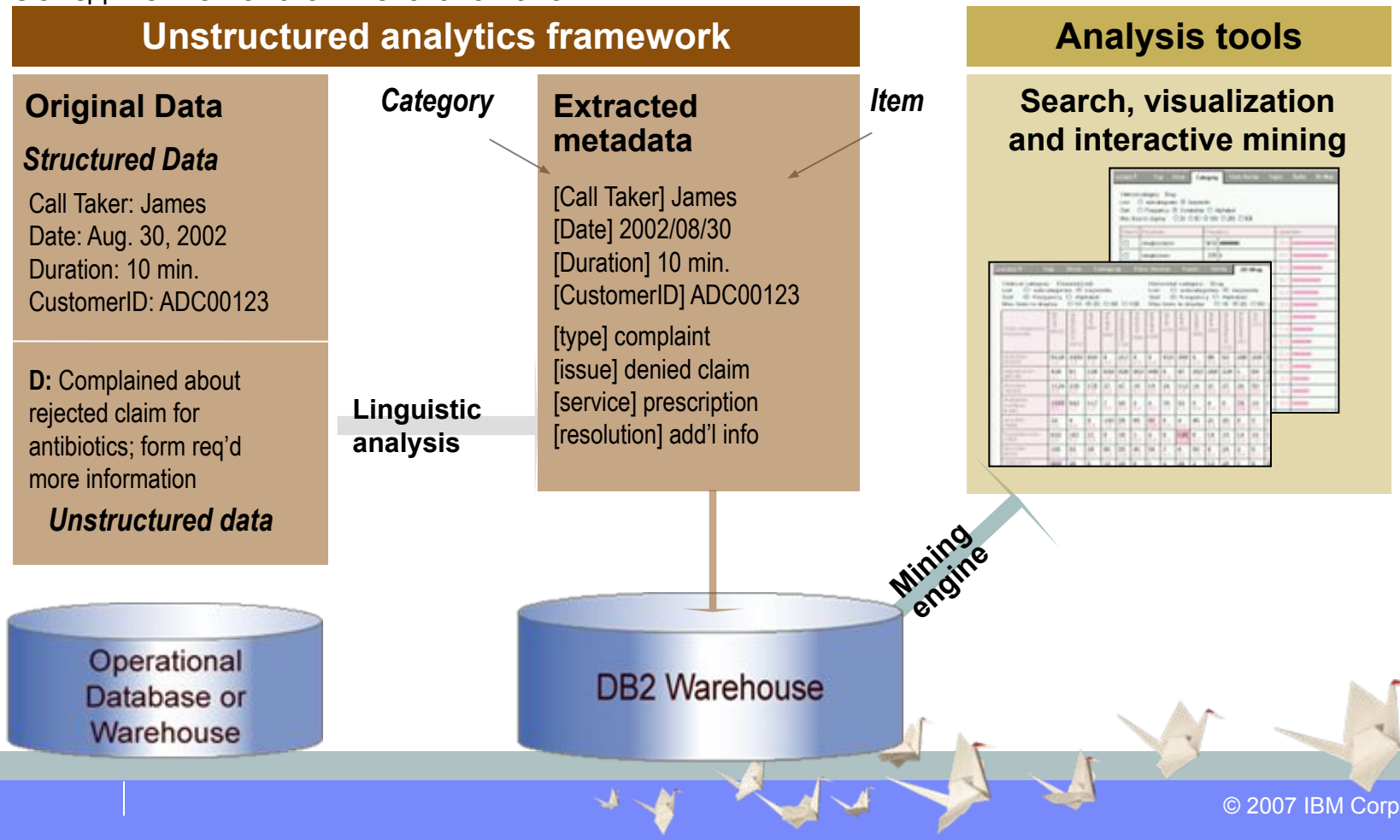


© 2007 IBM Corporation

Extended Insight

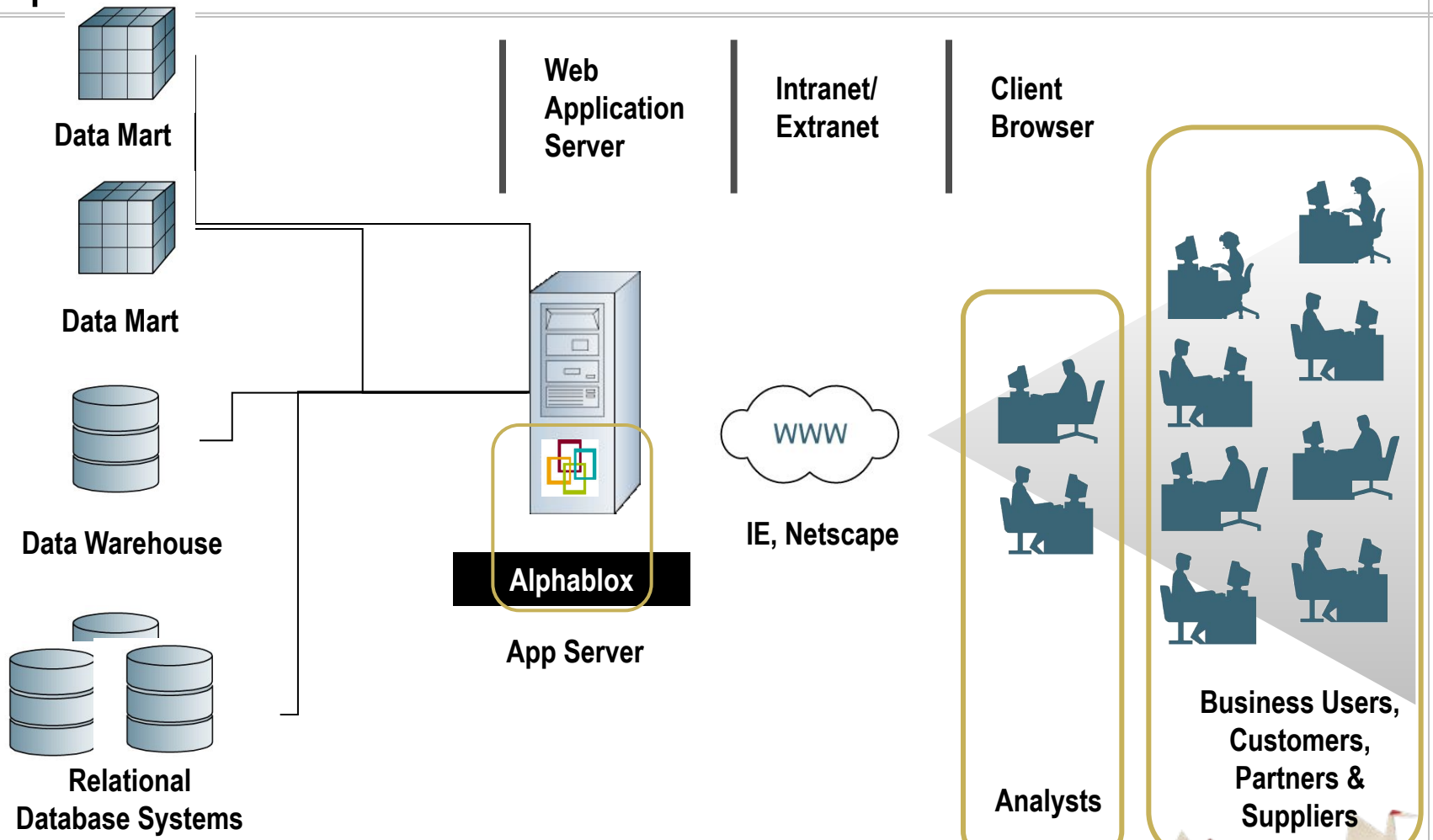
Introducing IBM OmniFind Analytics Edition

- Обширные интерфейсы для объединения структурированных и не структурированных данных
- Объединение поиска и текстовой аналитики



© 2007 IBM Corporation

Alphablox



© 2007 IBM Corporation

AlphaFinancial - DEMO - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Search Favorites Media

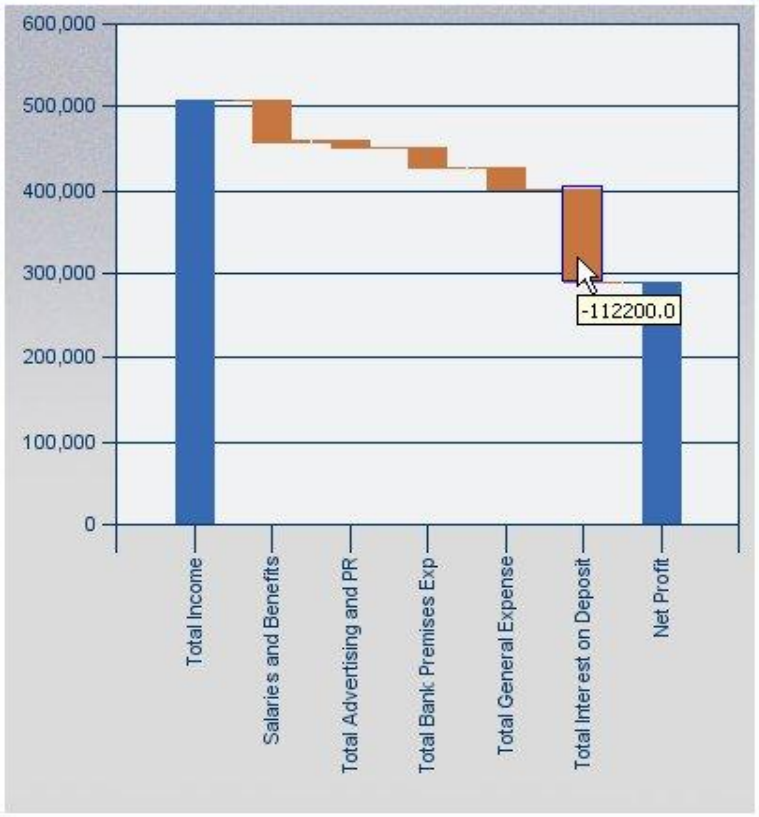
Address http://localhost:9080/AlphaFinancial2/start.jsp?savedstate=autosave&render=dhtml Go Links DB2 Alphablox 5.6 WebSphere Admin Console

alphafinancial

Welcome Betty!

- Home
- Branch View
  - Financials
    - Income Statement
  - Metrics
    - New Accounts
    - New Deposits
    - Ranking
  - Budget
    - Adjust Budget

	October		
	Actual	Budget	Variance
ees	$61,000	$80,000	$19,000
	$64,000	$80,000	$16,000
ome	$49,000	$80,000	$31,000
	$85,000	$90,000	$5,000
	$76,500	$80,000	$3,500
	$60,000	$105,000	$45,000
	$111,500	$130,000	$18,500
	$507,000	$645,000	$138,000
	$48,000	$80,000	$32,000
	$5,800	$55,100	$49,300
	$24,000	$110,000	$86,000
	$27,900	$70,000	$42,100
ng	$55,000	$45,000	-$10,000
g	$56,000	$45,000	-$11,000
	$400	$800	$400
	$800	$1,200	$400
	$112,200	$92,000	-$20,200
	$217,900	$407,100	$189,200
	$289,100	$237,900	-$51,200
	$1,019,000	$1,290,000	$271,000



User Details

Country: US
Region: West
State: CA
City: San Francisco
Branch: 012
User: Branch, Betty A

Logout

AF News

9:02am EST

AF Reports - Rev Increases 76% - WW Trades Grow 62%

NEW! You have 1 NSF items for review. Go...

Done Local intranet

Start \\.host\Shared Folders\M... Paint Shop Pro - Image27 AlphaFinancial - DEMO ... 10:35 PM

Thank You

© 2007 IBM Corporation