

# X5 RETAIL GROUP

Развитие корпоративного  
хранилища данных в  
X5 Retail Group



## О компании

X5 Retail Group — ведущая продуктовая розничная компания России. Управляет портфелем брендов сетевых магазинов «Пятёрочка», «Перекрёсток» и «Карусель».



# Миссия X5

Обеспечить каждому россиянину, вне зависимости от его места жительства и достатка, возможность покупать продукты в современных магазинах, гарантирующих комфортный доступ к широкому ассортименту качественных товаров.



# Цель X5 Retail Group —

**быть самой ценной компанией** в глазах покупателей, сотрудников и акционеров, что означает:



## Признание покупателей

лидерство по росту выручки,  
показателя NPS и LFL-трафика



## Признание сотрудников

высокая вовлеченность  
и низкая текучесть



## Признание акционеров и инвесторов

лидерство по ключевым  
показателям рыночной оценки  
компании

# Ключевые стратегические приоритеты X5



Совершенствование ценностного предложения и ассортимента под потребности покупателей



Развитие сервиса наивысшего качества, высокий показатель NPS



Сбалансированный рост



Совершенствование операционной эффективности



Развитие технологичных, инновационных решений и big data



Развитие омниканальности

# Ключевые цифры X5



>20 лет

на рынке продуктового ритейла в России



>11%

Доля рынка



17 025

Магазинов\*



>307 тыс.

Сотрудников



>1,7 трлн руб.

Выручка в 2019 году

492,9 млрд руб.

чистая розничная выручка\*



5,2 млрд

Покупателей в год

14,2 млн

Покупателей в день

42 млн

Лояльных покупателей

# Региональное присутствие X5

66

регионов присутствия

42

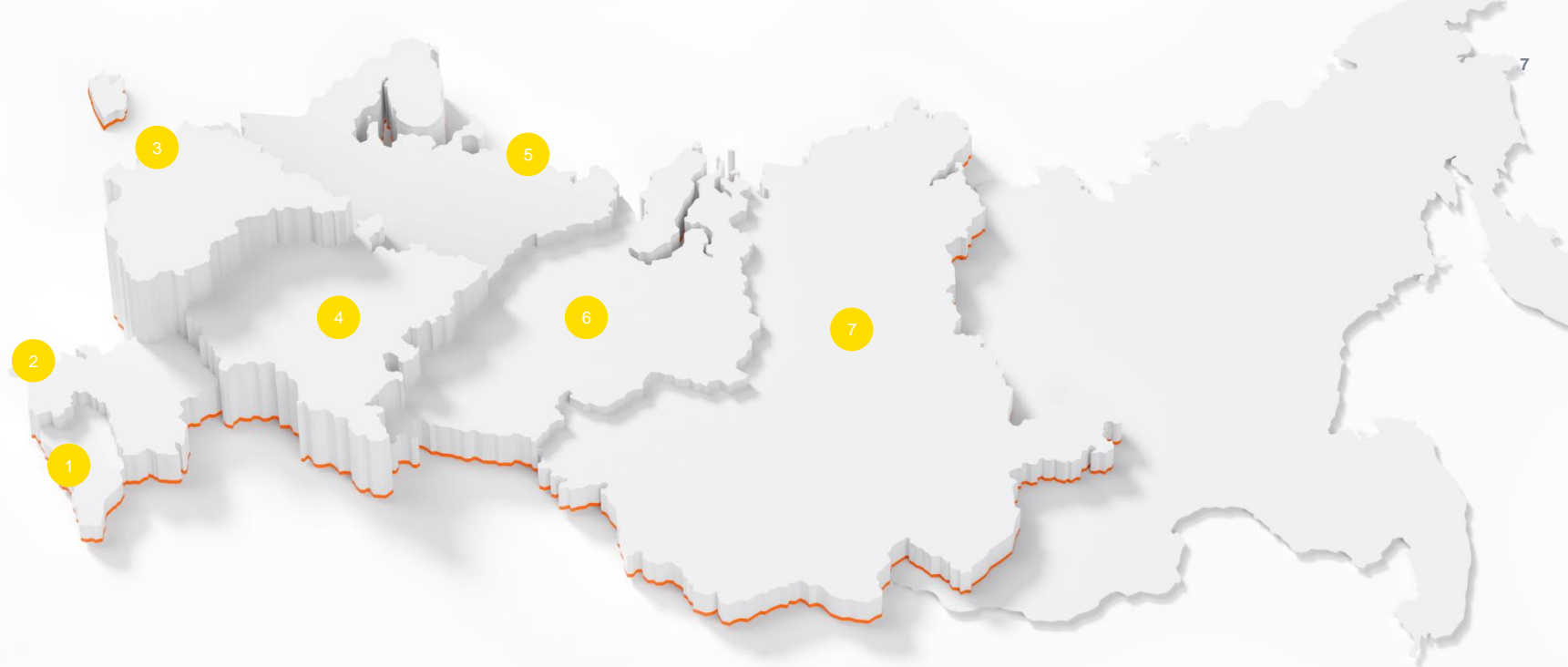
дистрибуционных центра

5900

поставщиков из России, СНГ, Европы, Азии, Африки, Латинской Америки

4083

собственных грузовых автомобиля



	скФО ①	юФО ②	цФО ③	пФО ④	сЗФО ⑤	уФО ⑥	сФО ⑦	ВСЕГО
Пятерочка	299	1542	6003	4335	1734	1392	791	16 096
Перекресток	9	51	506	130	121	50	-	867
Карусель	1	4	17	8	6	6	1	42
Количество РЦ	1,0%	6,1%	50,0%	18,8%	14,5%	6,8%	2,8%	100%
Чистая розничная выручка	1,0%	6,1%	50,0%	18,8%	14,5%	6,8%	2,8%	100%



# Roadmap (2009-2018)



## Объем проекта:

- Oracle 8
- BW 7.0

- Oracle 10
- Oracle 11g
- BW 7.3

- Business Warehouse Accelerator (ускорение формирования отчета ~5-10 раз)

- Миграция с 120 Tb Oracle в 12 Tb SAP HANA cloud
- SAP BW 7.4
- Ver 3.5 → 7.0
- Реализация чековой отчетности + COO
- NLS
- Hana RT 12ТБ до 54 ТБ

## Предпосылки:

- Объединение разнородных ХД в единое КХД

- Повышение стабильности работы системы

- Медленная работа отчетности

- Долгая загрузка данных
- Необходимость актуализации данных в BWA
- Долгий доступ к холодным данным (архивация данных на ленту)



# SAP BW on HANA | Миграция с Oracle on primes на HANA cloud #2

## Oracle



### ETL

Основные загрузки + загрузки BWA



## 120 TB + 4 TB

Oracle + BWA



## 2 500

Активных пользователей



## 20с

Среднее время формирования отчета

## SAP HANA



### ETL

Ускорение предоставления данных



## 55 TB

HANA. InMemory



## 3 500

Активных пользователей



## 9с

Среднее время формирования отчета

# Roadmap (2019-2021)

2019-2020(I, II)

2020(III , IV) - 2021

## Объем проекта:

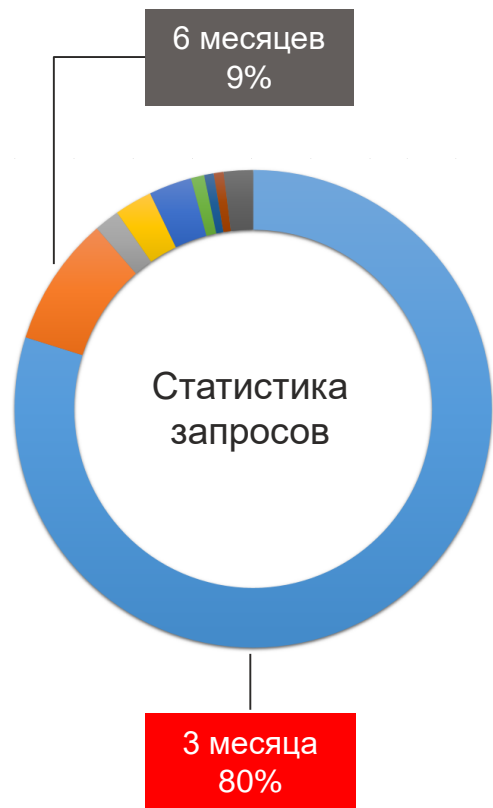
- HANA RT → HANA EE
- Hana EE (55Tb → 22 Tb)
- Вынос чековой функциональности на Greenplum (COO в Hana)
- Разработка интеграции HANA с Hadoop
- Охлаждение данных NLS + Hadoop
- Перенос Blue Book с XLS на модель HANA + Tableau

- SAP HANA 2.0 / SAP BW 7.5
- SAP Data Hub
- SAP DS
- SAP SDI
- SAP SLT
- Выбор базы данных для SSA

## Предпосылки:

- Высокая стоимость владения SAP HANA
- Отсутствие доступа к данным из сторонних приложений
- Снижение нагрузки на HANA
- Проблема в работе NLS
- Агрегаты по чекам занимают большой объем данных
- Расширение возможностей обработки и анализа данных
- Увеличение объема и качества обрабатываемых данных
- Ускорить доставку данных до хранилища

# Охлаждение HANA | Целевой подход к мультитемпературному хранению



## SAP



## DMP

Критичные данные

Теплые и горячие данные

Холодные данные

SAP HANA

24M

4M

SAP Sybase IQ

24M

hadoop

24M+

24M+

24M+

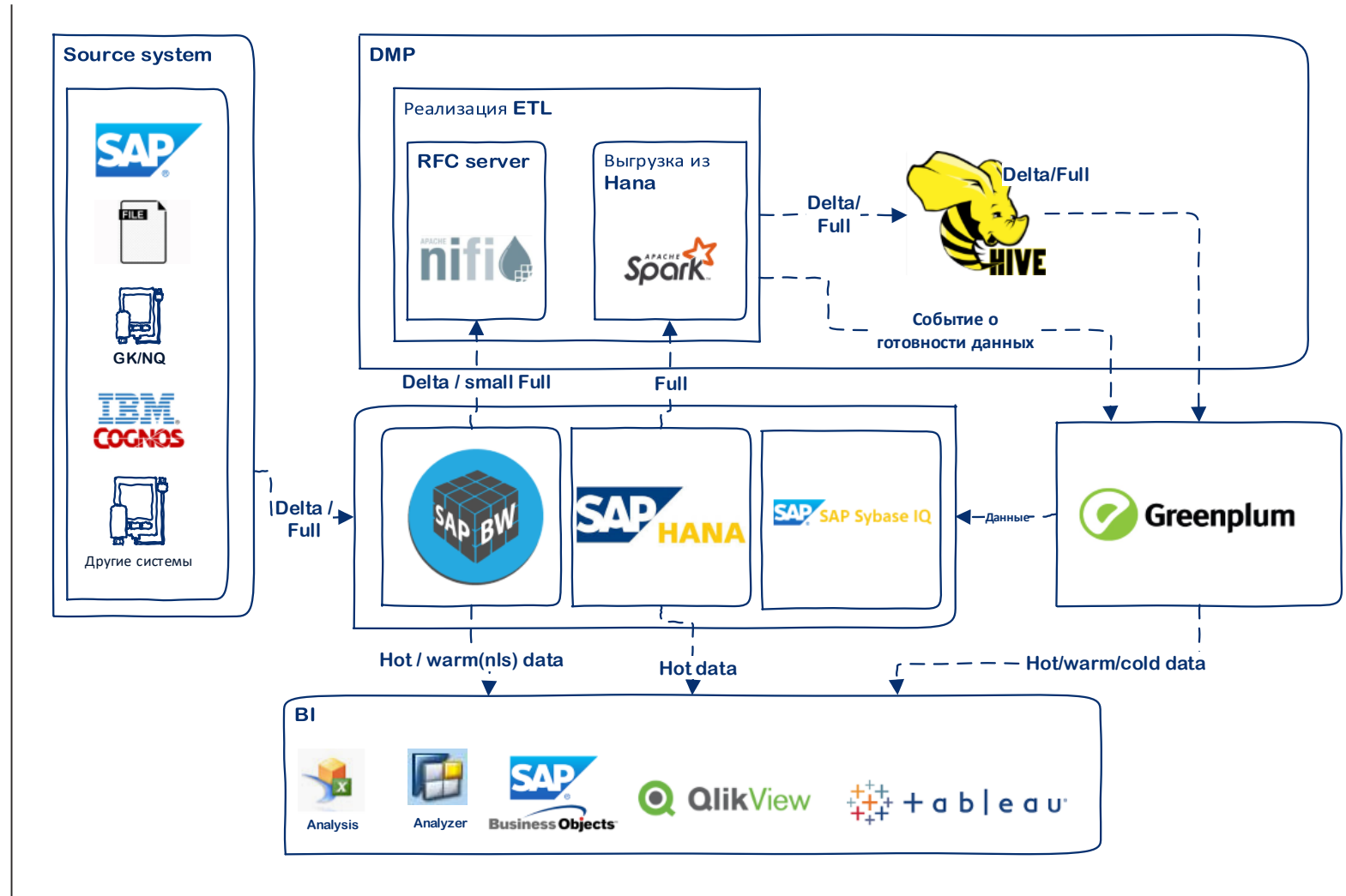
Greenplum

24M

24M

# Интеграция DMP & SAP BW | Целевая архитектура

- Преобразование схемы-звезды SAP BW в плоские таблицы
- SAP BW: ETL-инструмент
- HIVE: данные доступны для любых продуктов/задач
- Greenplum: данные доступны для построения отчетности
- Greenplum может быть заменен на любую реляционную БД
- Сохранение работоспособности ~95% отчетов



# Управление распределением данных SAP HANA scale-out

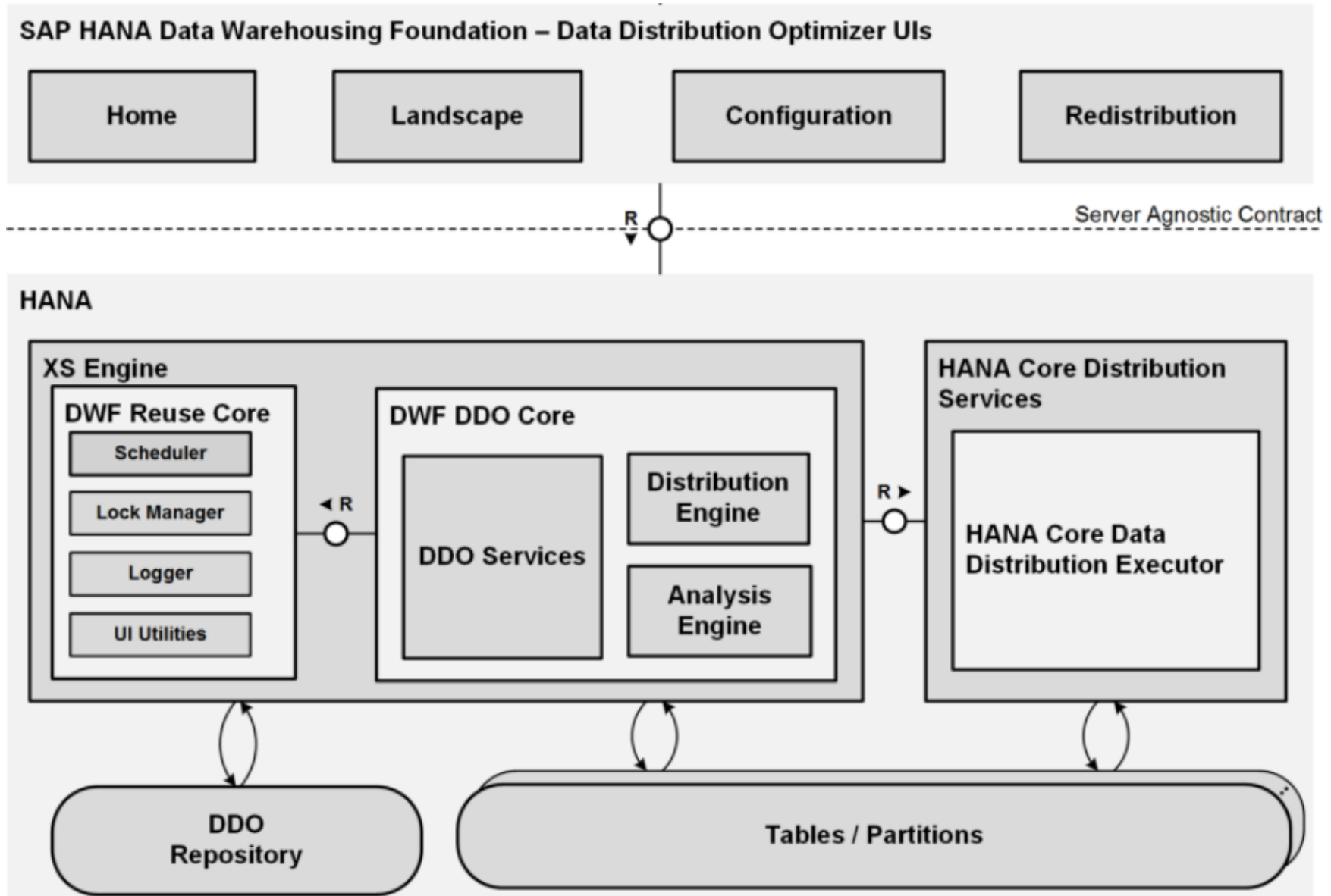
Для управления данными в Scale-out системе мы используем SAP HANA Data Warehousing (SAP HANA XS classic-based)

**Data Distribution Optimizer** основная цель помочь администратору SAP HANA управлять распределением таблиц и оптимизировать выделение памяти SAP HANA в горизонтально масштабируемой среде SAP HANA

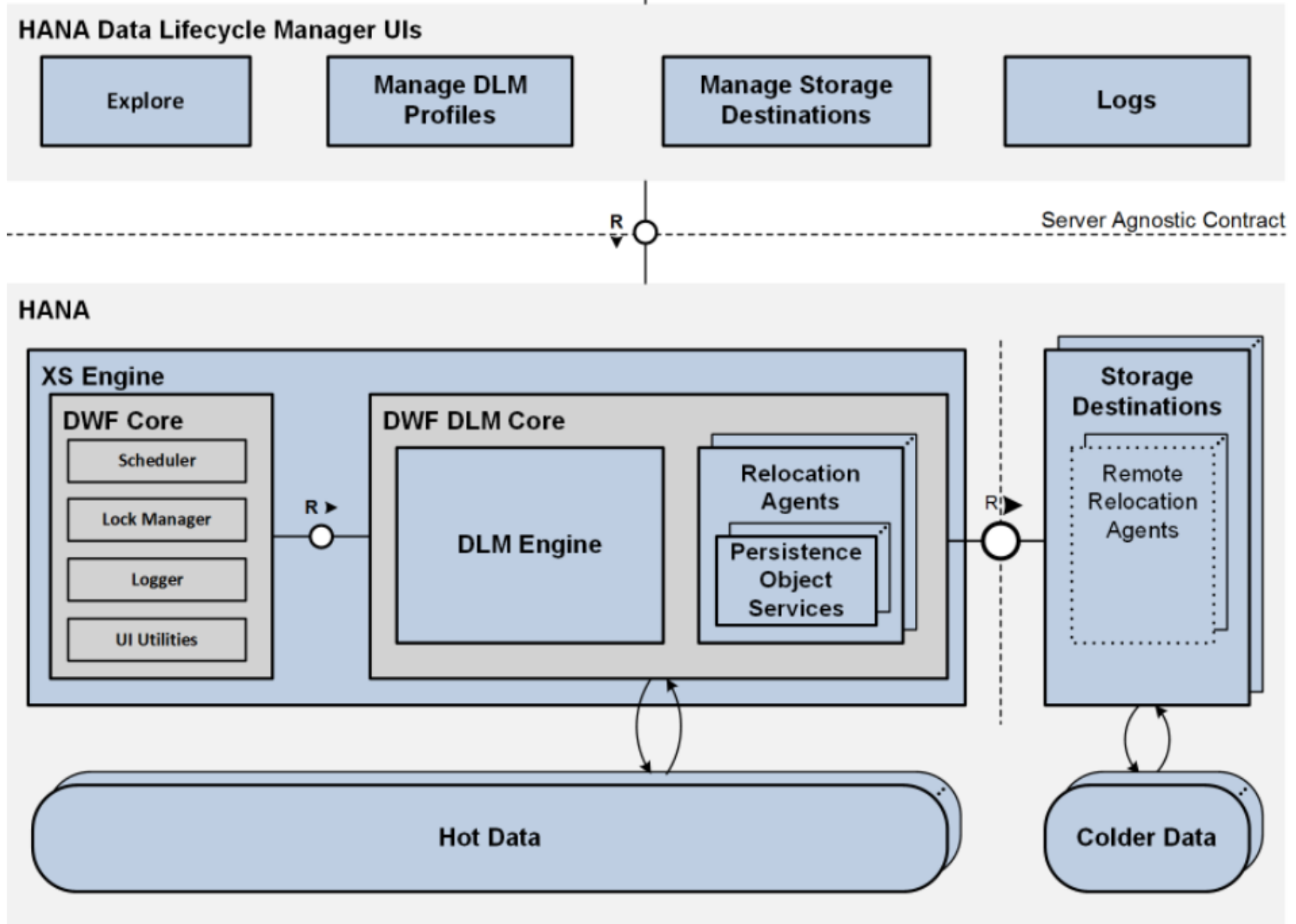
**Data Lifecycle Manager** представляет средства для удаления данных из SAP HANA persistency.

С его помощью вы можете использовать следующие хранилища: SAP HANA Dynamic Tiering, Hadoop (Spark SQL), SAP HANA extension node или SAP IQ in SAP HANA для хранения устаревших данных

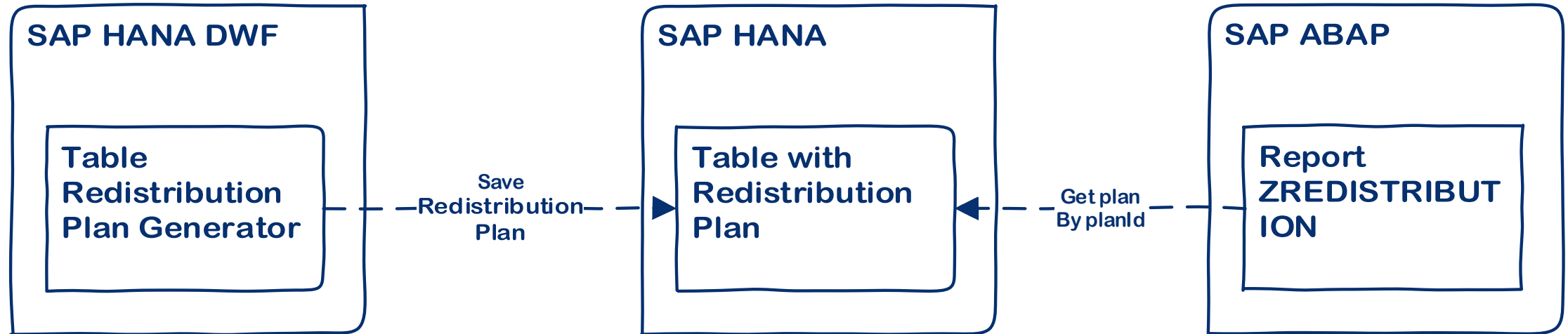
# Data Distribution Optimizer



# Data Lifecycle Manager



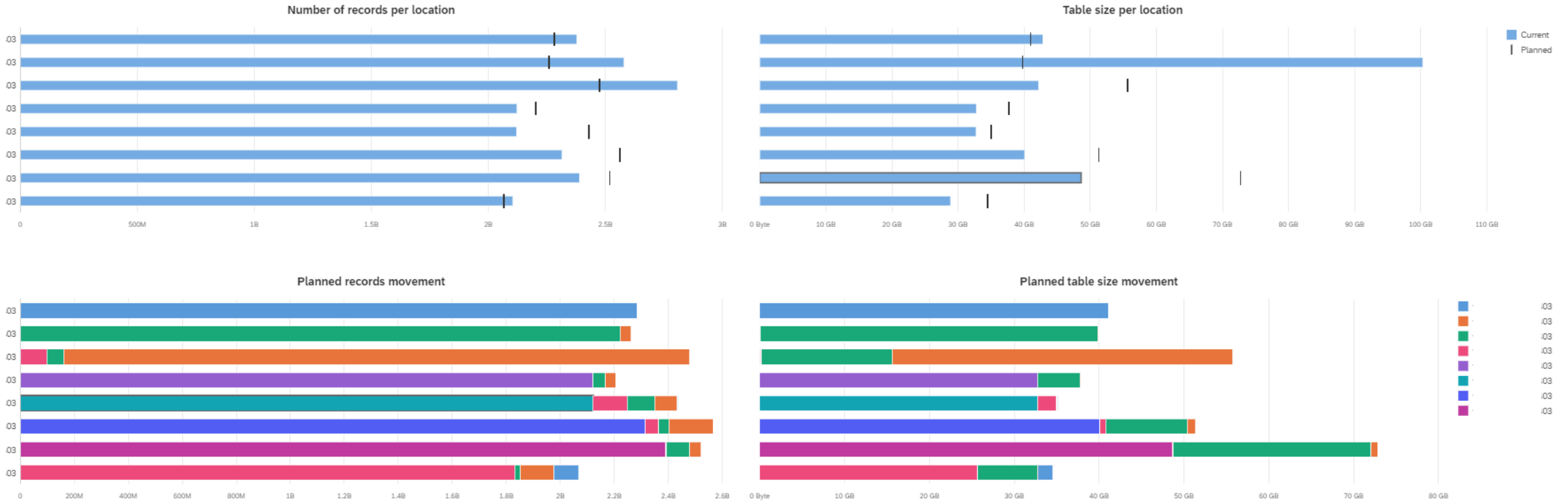
# Simulated with Actual System Landscape



- Управление нагрузкой и параллельностью выполнения редистрибуции
- Решена проблема 2 млрд таблиц/партиций
- Ускорение процедуры редистрибуции примерно в 4-е раза
- Полное планом редистрибуции
- Можно выполнять по частям
- Можно повторять упавшие шаги
- Настройка отдельных TablePlacement



# Анализ плана редистрибуции



# Автоматизация обновления данных в BI системах

