



**Российский университет дружбы народов
Институт гостиничного бизнеса и туризма**

В. Дихтяр

ИНФОРМАЦИОННЫЕ ТЕХНОЛОГИИ
(для бакалавров)

Раздел 1. Системы управления базами данных и базами знаний

Тема 1-4. Функциональные возможности СУБД

Самым важным изобретением, которое в будущем увидит свет в корпоративной исследовательской лаборатории, станет сама корпорация.

Поэтому необходимо разрабатывать не просто новые продукты, а новые технологические и организационные структуры, которые позволят создать постоянно обновляющуюся фирму.

Б.Геймс

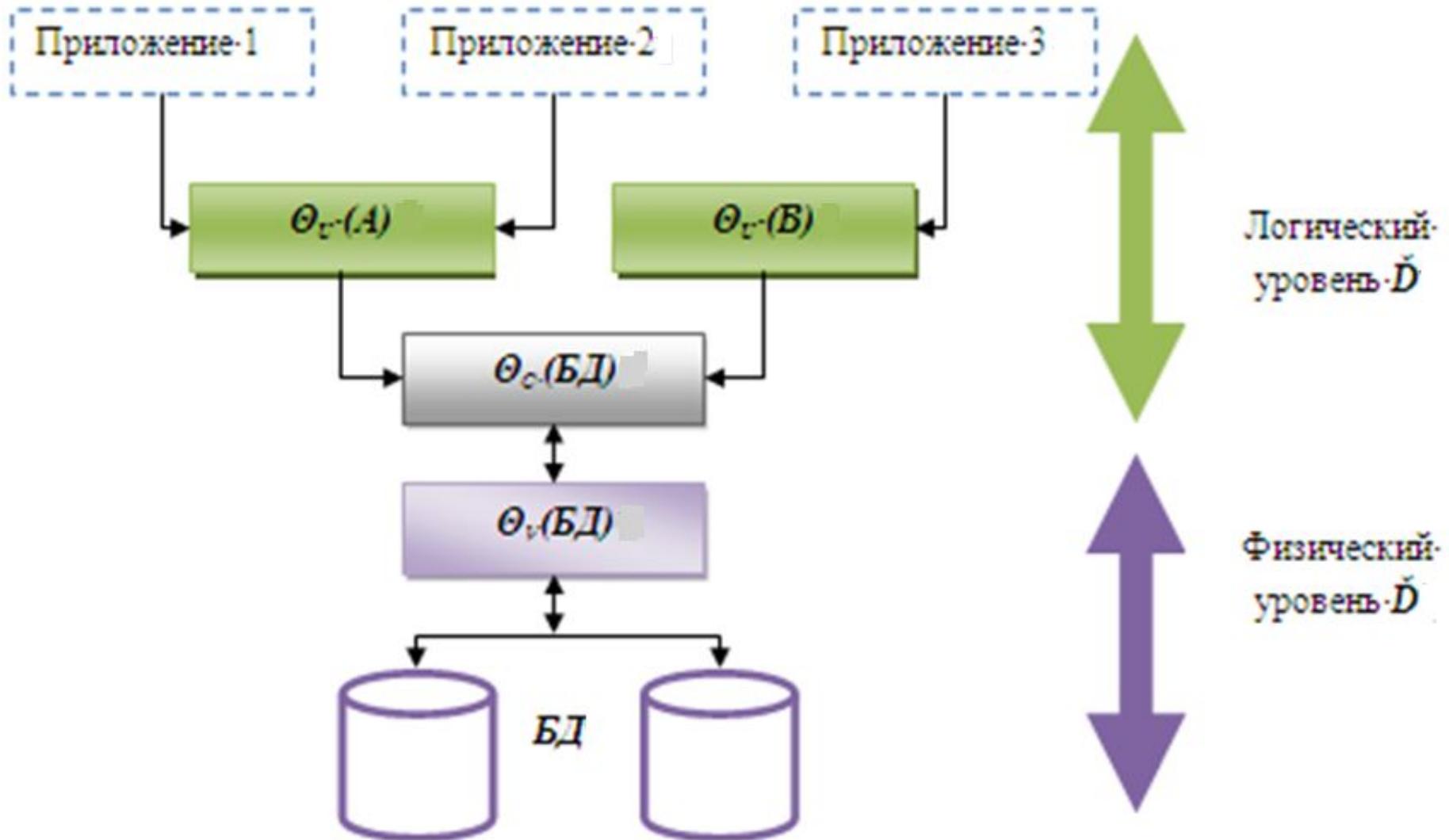
СОДЕРЖАНИЕ

- Архитектура СУБД
- Уровни представления данных и модели
- Информационно-логическая модель предметной области
- СУБД: классы, производительность, факторы влияния

УРОВНИ ПРЕДСТАВЛЕНИЯ $\check{D} \Rightarrow \Theta$ T 1.4

- Θ_C Концептуальная модель \equiv {экземпляры типов \check{D} , структурированы по требованиям СУБД} (*логический аспект*)
- Θ_V Внутренняя модель \equiv {экземпляры записей, физически хранимых в U -носителях} (*физический аспект*)
- Θ_U Внешняя модель $\equiv \subset \Theta_C \Rightarrow$ логическая $\hat{S}(\check{D})$ для приложения или ξ соответствует Θ_U (*частная $\hat{S}(\check{D})$ для ξ*)

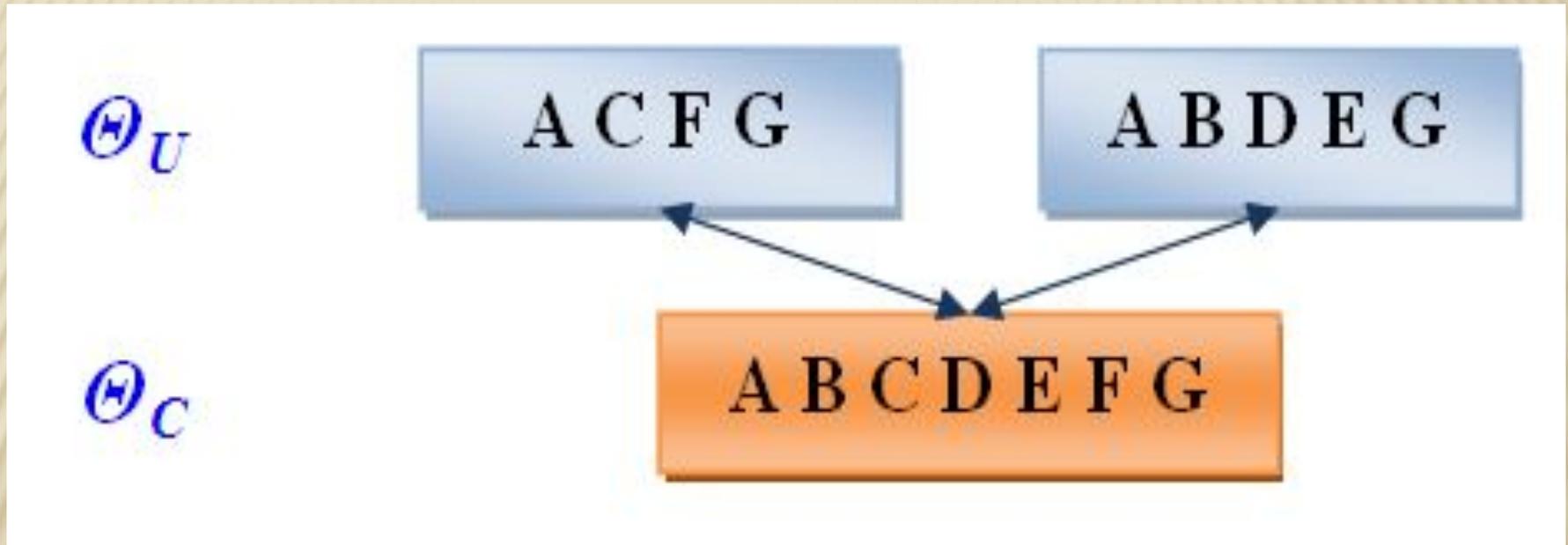
АРХИТЕКТУРА СУБД



Θ_U

- поддержка доступа к \check{D} (БД) приложений
- ограничен состав и $\hat{S}(\check{D}) \Theta_C$, доступных в приложении
- заданы допустимые режимы обработки \check{D} : ввод, редактирование, удаление, поиск

□ СООТНОШЕНИЕ $\Theta_C \leftrightarrow \Theta_U$ □



Θ_C - концептуальная модель , Θ_U - внешняя модель

$\Delta\Theta$

- $\Delta(\check{\text{т}}\text{-потребностей}) \Rightarrow \Delta\Theta_U$, **не** $\Delta\Theta_V$ и Θ_C
- $\Delta\Theta_C$: \perp виды \check{D} или $\Delta\hat{S}(\check{D})$ могут затрагивать не все приложения \Rightarrow независимость $n_{\check{D}}$ от \check{D}
- $\Delta\Theta_C \Rightarrow \Delta\Theta_V$
- Θ_C не изменяется \Rightarrow возможно самостоятельное $\Delta\Theta_V$ для \uparrow характеристик (расход памяти U -устройств..)

\Rightarrow **принцип** относительной независимости логической и физической организации \check{D}

Θ - моделирование, Θ_U Внешняя модель, Θ_V Внутренняя модель, Θ_C Концептуальная модель, \check{D} - данные, \check{S} - систем/а –ный,

$\Theta(\Omega)$

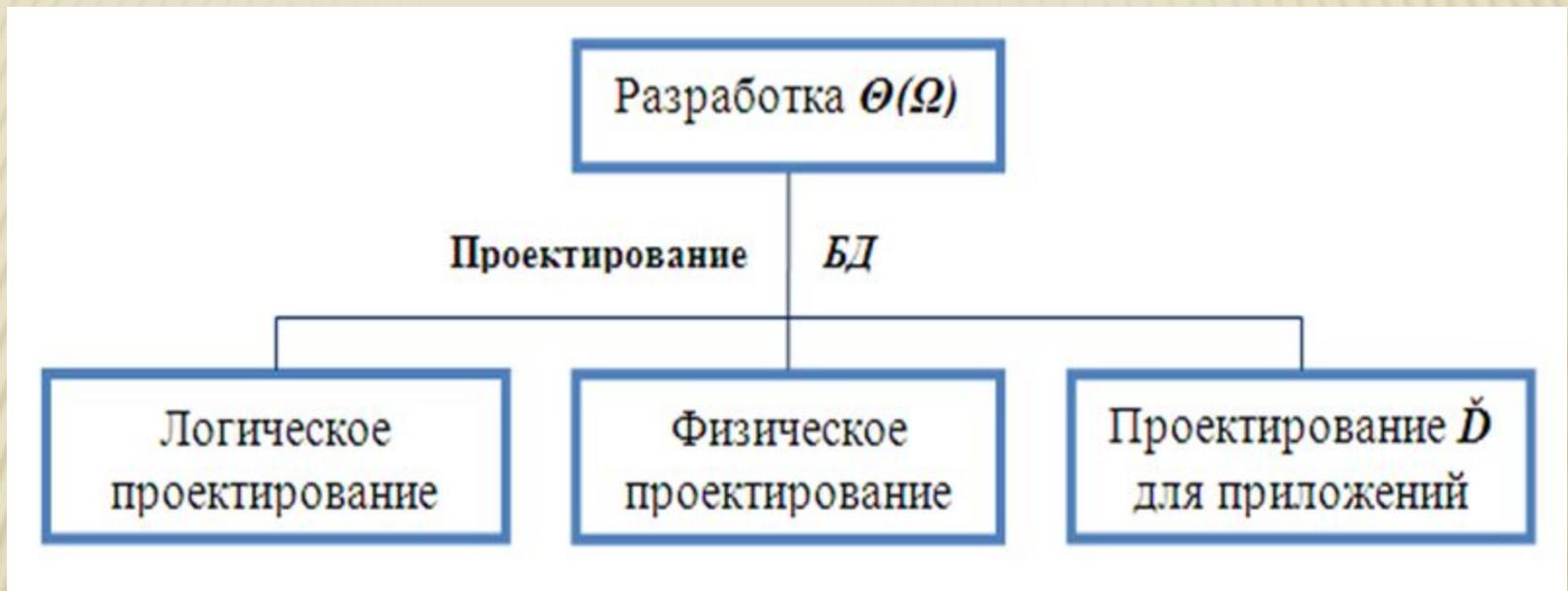
$\Theta(\Omega) \equiv$ Информационно-логическая модель предметной области \equiv отражение $\Omega: \{\hat{O} + r\}$

в интегрированном виде отражаются:

- состав и $\hat{S}(\check{D})$
- \check{i} -потребности приложений (задач и запросов)

$\Theta(\Omega)$ - Информационно-логическая модель предметной области, Ω - предметная область, \check{D} - данные, \check{S} - систем/а –ный

ЭТАПЫ ПРОЕКТИРОВАНИЯ БД



- предварительная Θ
- уточнение
- Θ_C □ Θ_V □ Θ_U

Θ - моделирование Θ_U Внешняя модель, Θ_V Внутренняя модель, Θ_C Концептуальная модель

□ ГРАФИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ $\Theta(\Omega)$ □



$\Theta(\Omega)$ - Информационно-логическая модель предметной области

СУБД

≡ *п*-система для

- создания *БД* для $\{\check{R}(\zeta)\}$
- поддержания *БД* в актуальном состоянии
- обеспечения доступа ξ к \check{D} в рамках полномочий
- централизованного управления *БД*

п- программа, \check{D} - данные

КЛАССЫ СУБД

Общего назначения \equiv \mathfrak{P} -комплексы: { f создания и эксплуатации $\mathcal{B}\mathcal{D}$ }

- не ориентированы на конкретные \mathcal{Q}
- реализуются как $\mathfrak{P}\dot{h}$ и коммерческое изделие
- обладают средствами настройки на конкретную $\mathcal{B}\mathcal{D}$
- позволяют \downarrow (t (разработки), трудовые \dot{R})
- имеют развитые f -возможности и даже избыточность

Специализированные : при невозможности / нецелесообразности СУБД общего назначения

\mathfrak{P} - программа, \mathcal{Q} - предметная область, \dot{h} - продукт

ПРОИЗВОДИТЕЛЬНОСТЬ СУБД (P)

- t выполнения запросов
- скорость поиска i
- t импортирования **БД** из других форматов
- скорость создания индексов и выполнения массовых операций (обновление, вставка..)
- число параллельных обращений к **Д** в многопользовательском режиме
- t генерации отчета

Д- данные

ФАКТОРЫ ВЛИЯНИЯ НА ПРОИЗВОДИТЕЛЬНОСТЬ СУБД (P_{DB})

- СУБД: соблюдение целостности $\check{D} \Rightarrow$
дополнительная нагрузка (не испытывают другие \check{P})
- β собственных прикладных \check{P} сильно зависит от
правильного проектирования и построения $БД$

\check{D} - данные, \check{P} - программа, β - производительность

ЦЕЛОСТНОСТЬ \check{D} НА УРОВНЕ БД

\check{I} в БД всегда остается корректной и полной \Rightarrow

- установить правила целостности (соблюдаться на глобальном уровне)
- обеспечивать независимо от способа ввода \check{D} (интерактивно, импорт, специальная \check{P})

\check{D} - данные, \check{P} - программа

СРЕДСТВА ОБЕСПЕЧЕНИЯ ЦЕЛОСТНОСТИ

- назначение первичного ключа

средства для работы с типом полей с автоматическим приращением (СУБД присваивает новое уникальное значение)

- поддержание ссылочной целостности:

запись \bar{I} о связях таблиц

пресечение любой операции, приводящей к нарушению ссылочной целостности

ЦИФРОВАЯ РЕВОЛЮЦИЯ

- сбор больших объемов **Д** с помощью устройств автоматического формирования первичных **Д**
- 8 **С** оплаты, штрих-коды, смарт-карты □
- **Д** м. б. сохранены в объемных, современных и не дорогих **С** хранения

Ў И ПРИНЯТИЕ Ў

Сами массивы \mathcal{D} не представляют ценности для \mathcal{m} :

- хранятся в разных форматах или ед.измерения;
- могут находиться в различных местах;
- \mathcal{D} операционных \mathcal{S} обработки транзакций (СОТ) отражают текущий статус, а для поддержки \mathcal{R} необходимы временные ряды

□ сведения с января по декабрь □

⇒ принять $\mathcal{R}_{\mathcal{m}}$: Q - и q -анализ + \mathcal{i} и \mathcal{Z}

\mathcal{D} - данные, \mathcal{S} - систем/а –ный, \mathcal{Z} - знания

СОЗДАНИЕ ХРАНИЛИЩ \check{D}

- \perp подход к управлению *БД*

связующее звено между ограниченными возможностями \check{S} обработки транзакций и потребностями в достоверной \check{I} и \check{Z}

Хранилища ($\check{D}W$) \equiv базовая архитектура для *СППР* и управляющих *ИС*

ζ : обеспечить \check{m} своевременной и достоверной \check{I} и \check{Z} о Φ , $\check{h}(\Phi)$, ξ

\Rightarrow помочь эффективно управлять Φ

ДW

≡ централизованный архив реорганизованных **Д**:

- предметно-ориентированных
- интегрированных
- расположенных с соблюдением хронологии
- неизменных
- являющихся основой для поддержки принятия **Р**

≡ многомерная **БД**, отделенная от операционных баз

ОСНОВА ОРГАНИЗАЦИИ \check{D} В $\check{D}W$

- реляционные Θ (БД): \check{D} в виде двумерных таблиц
 - некоторые $\check{D}W$ используют все Θ (БД):
реляционные, иерархические и многомерные
 - многомерные базы:
 - \check{D} в виде n -мерного куба
 - ξ имеет дело с «сечениями» \check{D}
- по \dot{h} , районам, t .. + \uparrow выполнение запросов □

ПРЕДМЕТНАЯ ОРИЕНТИРОВАННОСТЬ \check{D}

- \check{D} организованы по предметам, или входам
 - ξ, A , оптовые Φ, h ..□
- G : обеспечить \check{I} по выбранному предмету
 - \check{S}_o : поддержка специальных f или видов A – вкладов, банковских карт, кредитов..□

ИНТЕГРИРОВАННОСТЬ \mathcal{D}

- использование денормализованных таблиц
- очистка \mathcal{D} от ошибок и избыточной \mathcal{I}
- добавление новых полей и ключей.
- сортировка, комбинирование и представление ξ

СОГЛАСОВАНИЕ Ъ

- по кодировкам
 - пол задается в виде «M,F», а не «смесью» обозначений типа «m,f», «Male,Female» □
- по единицам измерения
 - см или дюймы, не одновременно! □.

ХРОНОЛОГИЯ \checkmark

- \checkmark представляют собой «моментальные копии» и соответствуют ситуациям в различные T

НЕИЗМЕНЕННОСТЬ \checkmark

- не перезаписываются
- только для чтения, нельзя изменить
- статичны (операционные \checkmark непрерывно обновляются)

ВИТРИНЫ \check{D}

- Использование $\check{D}W$: создание и управление требуют T , усилия и M
- Некоторые Φ выбирают небольшие $\{\check{D}\}$, удовлетворяющих узкие потребности небольших групп ξ внутри Φ

Витрины $\check{D} \equiv$ небольшое, с ограниченным числом входов подмножество хранилища, предназначенное для небольшого сегмента ξ

□ ПРИМЕР СОЗДАНИЯ ВИТРИН □

Φ : 4 региональных отделения

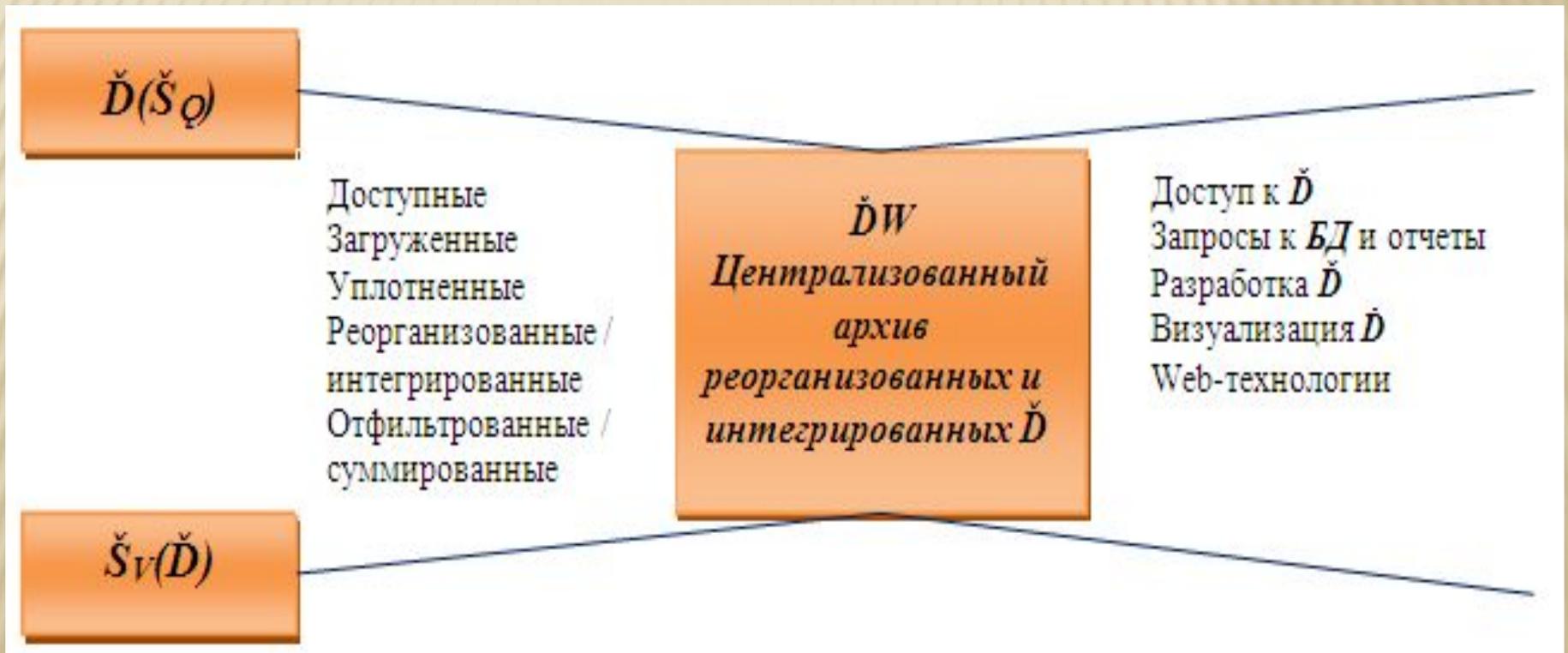
↓ q (запросов) \forall отделения о другом

\check{R} : 4 витрины вместо 1 хранилища

⇒ ↑ скорость $\times 4$: объем **\check{D}** для одного запроса ↓ на $\frac{3}{4}$

ПРОЦЕССЫ ОБРАБОТКИ Д В ДВ

Построение → Управление → Эксплуатация



СЛОВАРЬ СИМВОЛОВ

- ▣ **Д**- данные
- ▣ **З**- знания
- ▣ **С**- систем/а –ный
- ▣ **Р**- производительность
- ▣ **П**- программа
- **Ω**- предметная область
- **h**- продукт
- **Θ**- моделирование