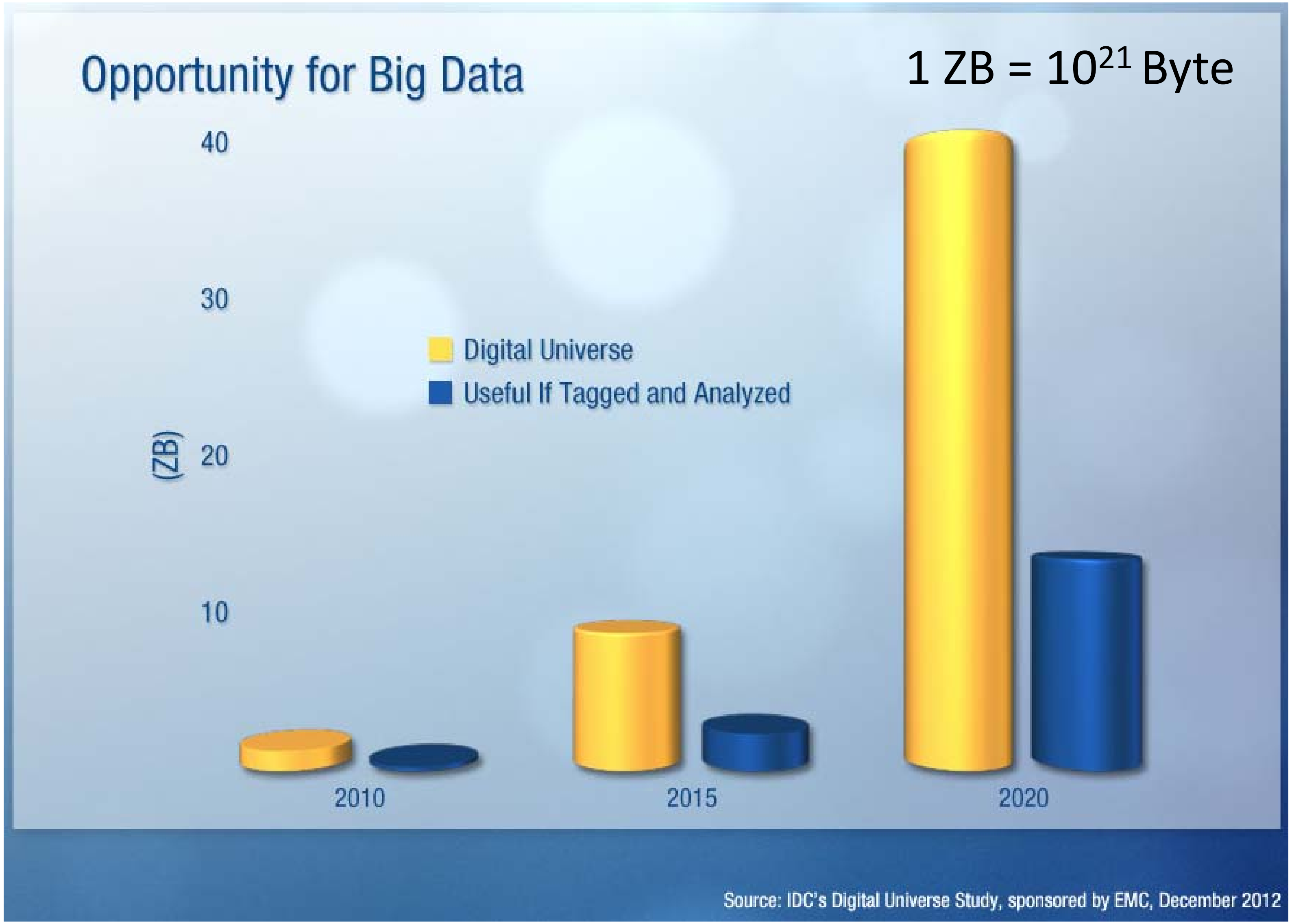
АКТУАЛЬНЫЕ ПРОБЛЕМЫ ВЫЧИСЛИТЕЛЬНОЙ
И ПРИКЛАДНОЙ МАТЕМАТИКИ – 2014

Использование распределенных колоночных индексов для выполнения запросов к сверхбольшим базам данных

Л.Б. Соколинский, Е.В. Иванова

Южно-Уральский государственный национальный
исследовательский университет

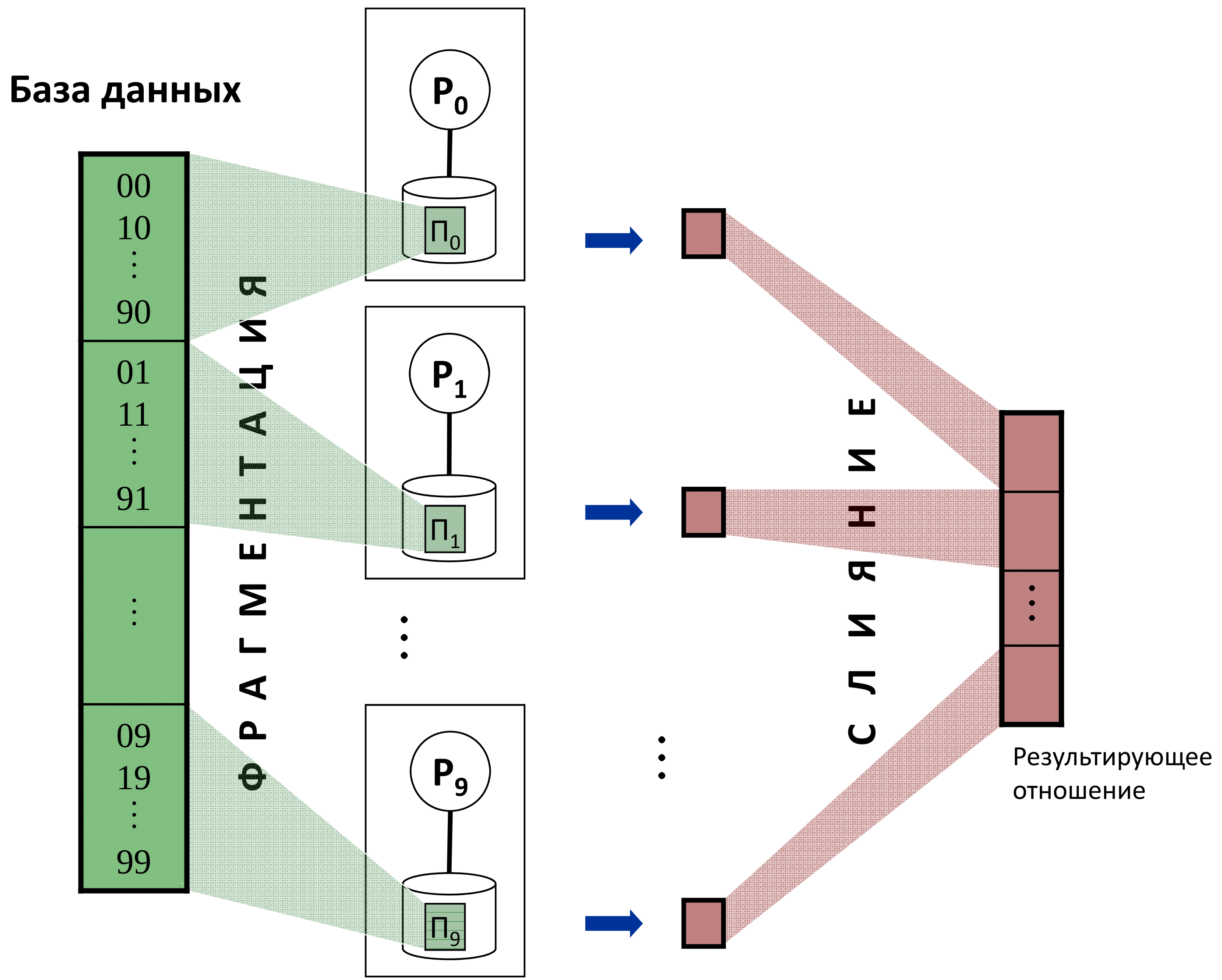
Экспоненциальный рост цифровой вселенной



Tianhe-2 (MilkyWay-2) - TH-IVB-FEP Cluster, Intel Xeon E5-2692 12C 2.200GHz, TH Express-2, Intel Xeon Phi 31S1P

Site:	National Super Computer Center in Guangzhou
Manufacturer:	NUDT
Cores:	3,120,000
Linpack Performance (Rmax)	33,862.7 TFlop/s
Theoretical Peak (Rpeak)	54,902.4 TFlop/s
Power:	17,808.00 kW
Memory:	1,024,000 GB

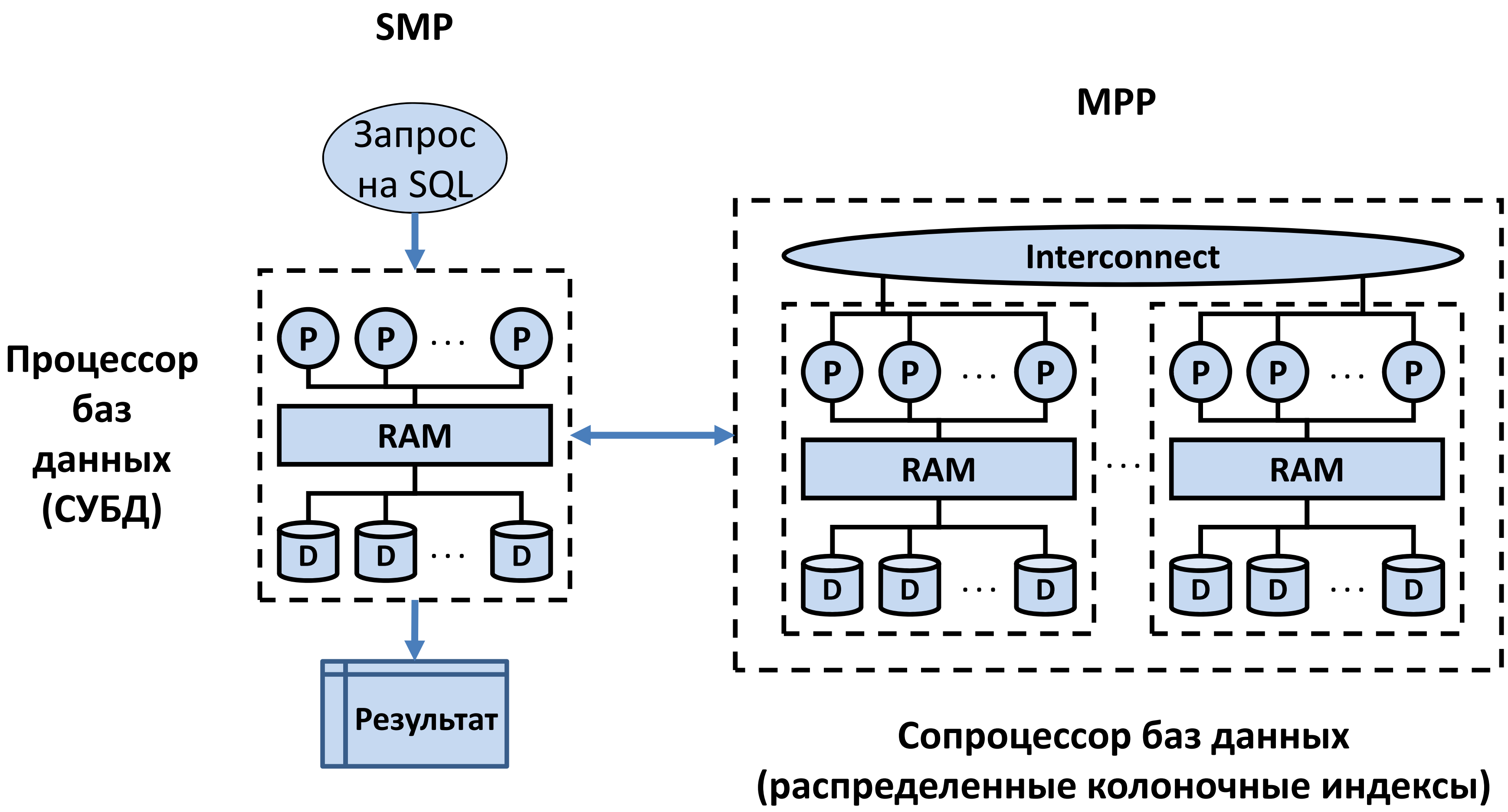
Общая схема параллельной обработки запроса



Почему трудно сделать эффективную СУБД для суперкомпьютера?

- Передача данных по сети
- Балансировка загрузки
- Низкая вычислительная сложность реляционных операций

Новая архитектура СУБД



Колоночный индекс

- Пусть $R(A^*, B, \dots)$– отношение R с первичным ключом A и атрибутом B. A представляет собой *служебный первичный ключ* и состоит из целочисленных неотрицательных элементов.
- $\mathfrak{D}_B$ – *домен атрибута* B. На множестве $\mathfrak{D}_B$задано отношение линейного порядка. $T(R) = n$ – количество элементов в R.
- ***Колоночным индексом*** $I_{R.B}$ атрибута B отношения R называется упорядоченное отношение, удовлетворяющее следующим требованиям:

$T(I_{R.B}) = n$ и $\pi_A(I_{R.B}) = \pi_A(R)$;

$\forall x_1, x_2 \in I_{R.B}(x_1 \le x_2 \Leftrightarrow x_1.B \le x_2.B)$;

$\forall r \in R(\forall x \in I_{R.B}(r.A = x.A \Rightarrow r.B = x.B))$.

R

A	B	....
0	36	....
1	14	....
2	36	....
3	10	....
4	74	....
5	27	....
6	58	....

$I_{R.B}$

A	B
3	10
1	14
5	27
2	36
0	36
6	58
4	74

Доменно-интервальная фрагментация

- Разобьем множество значений домена $\mathfrak{D}_B$ на k непересекающихся интервалов:

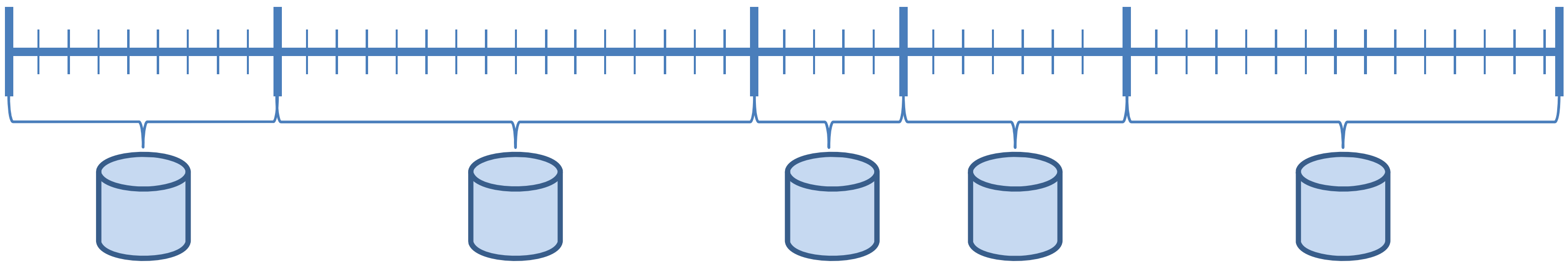
$$V_0 = [v_0; v_1]; V_1 = (v_1; v_2]; \ldots; V_{k-1} = (v_{k-1}; v_k];$$

$$v_0 < v_1 < \cdots < v_k;$$

$$\mathfrak{D}_B = \bigcup_{i=0}^{k-1} V_i$$

- ***Доменная функция фрагментации:***

$$\varphi_{\mathfrak{D}_B} : \mathfrak{D}_B \to \{0, \ldots, k-1\}$$



Доменно-интервальная фрагментация

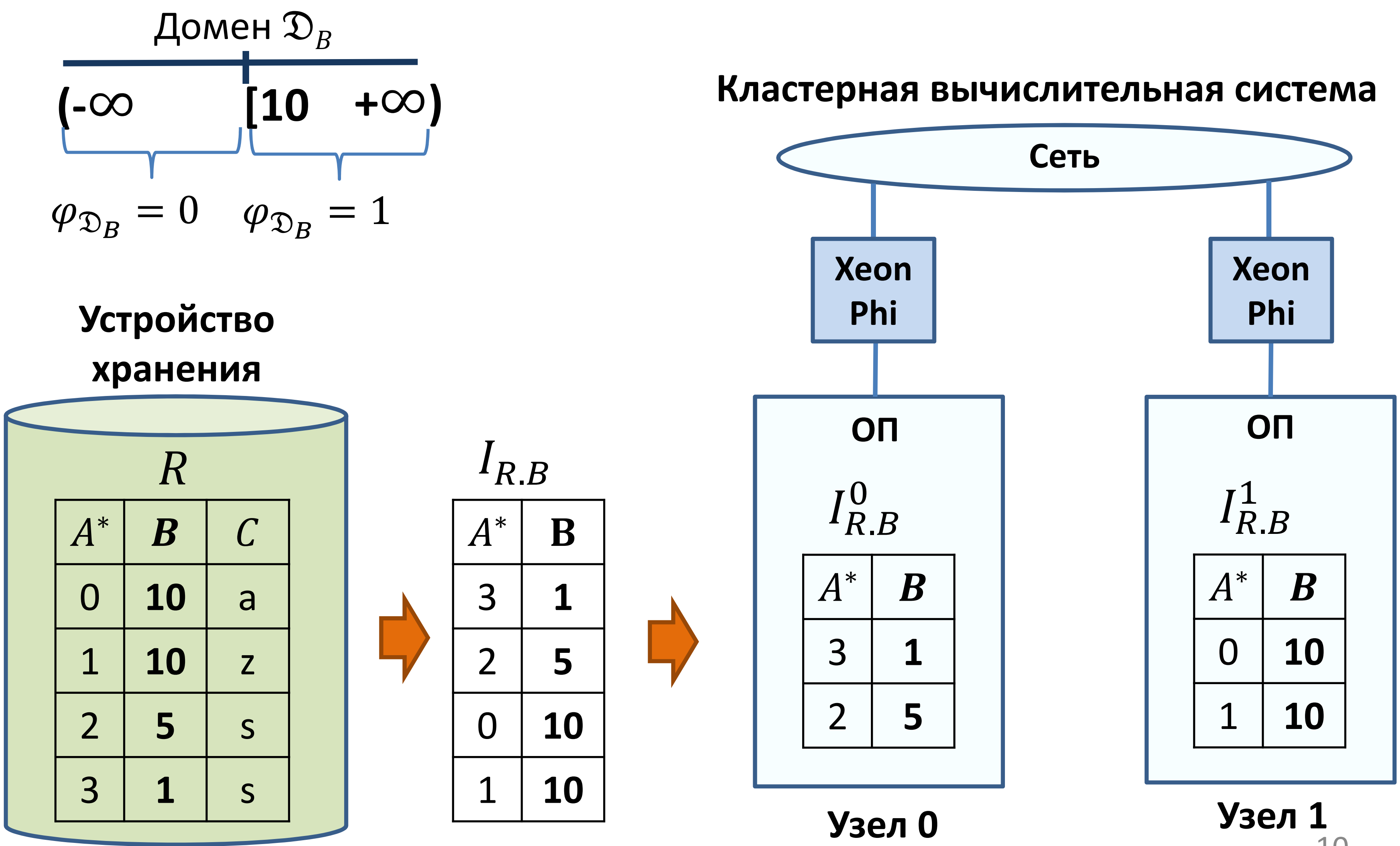
- ***Функция фрагментации индекса***:

$$\varphi_{I_{R.B}} : I_{R.B} \rightarrow \{0, \dots, k-1\}$$

$$\forall x \in I_{R.B} \left(\varphi_{I_{R.B}}(x) = \varphi_{\mathfrak{D}_B}(x.B) \right)$$

- Определим ***i-тый фрагмент индекса*** следующим образом $I^i_{R.B} = \left\{ x \mid x \in I_{R.B};\ \varphi_{I_{R.B}}(x) = i \right\}$

Пример доменно-интервальной фрагментации колоночного индекса



R

A^*	B	C
0	10	a
1	10	z
2	5	s
3	1	s

$I_{R.B}$

A^*	B
3	1
2	5
0	10
1	10

$I_{R.B}^0$

A^*	B
3	1
2	5

$I_{R.B}^1$

A^*	B
0	10
1	10

Операция естественного соединения

- Пусть даны $R(A^*, B_1, \dots, B_u, C_1, \dots, C_h)$ и $S(A^*, B_1, \dots, B_u, D_1, \dots, D_w)$.
- Определим Q как результат операции $R \bowtie S$.
- Пусть имеется два набора колоночных индексов по атрибутам $B_1, \dots, B_u$, для которых задана доменно-интервальная фрагментация степени k: $$I_{R.B_j} = \bigcup_{i=0}^{k-1} I^i_{R.B_j}; \quad I_{S.B_j} = \bigcup_{i=0}^{k-1} I^i_{S.B_j}.$$
- Положим для всех $i = 0, \dots, k-1$ и $j = 1, \dots, u$

$$P_j^i = \pi_{I^i_{R.B_j}.A \to A_R,\, I^i_{S.B_j}.A \to A_S} \left(I^i_{R.B_j} \underset{I^i_{R.B_j}.B_j = I^i_{S.B_j}.B_j}{\bowtie} I^i_{S.B_j} \right)$$

$$P_j = \bigcup_{i=0}^{k-1} P_j^i \qquad P = \bigcap_{j=1}^{u} P_j$$

$$Q = \left\{ r \circ (s.D_1, \dots, s.D_w) \mid r \in R \wedge s \in S \wedge (r.A, s.A) \in P \right\}$$

Выполнение операции естественного соединения

Определим $R \bowtie S$

Степень фрагментации: 2

Пусть имеются распределенные колоночные индексы для атрибута соединения B таблиц R и S

R

A	B	C
0	10	Ni
1	12	Au
2	5	Pb
3	1	Ag

S

A	B	D
0	5	Pb
1	11	Pb
2	3	Ni
3	10	Fr
4	3	Ag

Узел 0

$I^0_{R.B}$

A	B
3	1
2	5

$I^0_{S.B}$

A	B
2	3
4	3
0	5

Узел 1

$I^1_{R.B}$

A	B
0	10
1	12

$I^1_{S.B}$

A	B
3	10
1	11

Выполнение операции естественного соединения

Узел 0

$I^0_{R.B}$

A	B
3	1
2	5

$I^0_{S.B}$

A	B
2	3
4	3
0	5

P^0

$R.A$	$S.A$
2	0

Узел 1

$I^1_{R.B}$

A	B
0	10
1	12

$I^1_{S.B}$

A	B
3	10
1	11

P^1

$R.A$	$S.A$
0	3

Выполнение операции естественного соединения

Узел 0

$I^{0}_{R.B}$

A	B
3	1
2	5

$I^{0}_{S.B}$

A	B
2	3
4	3
0	5

P^0

$R.A$	$S.A$
2	0

P^1

$R.A$	$S.A$
0	3

Узел 1

$I^{1}_{R.B}$

A	B
0	10
1	12

$I^{1}_{S.B}$

A	B
3	10
1	11

P^1

$R.A$	$S.A$
0	3

Выполнение операции естественного соединения

Узел 0

P^0

$R.A$	$S.A$
2	0

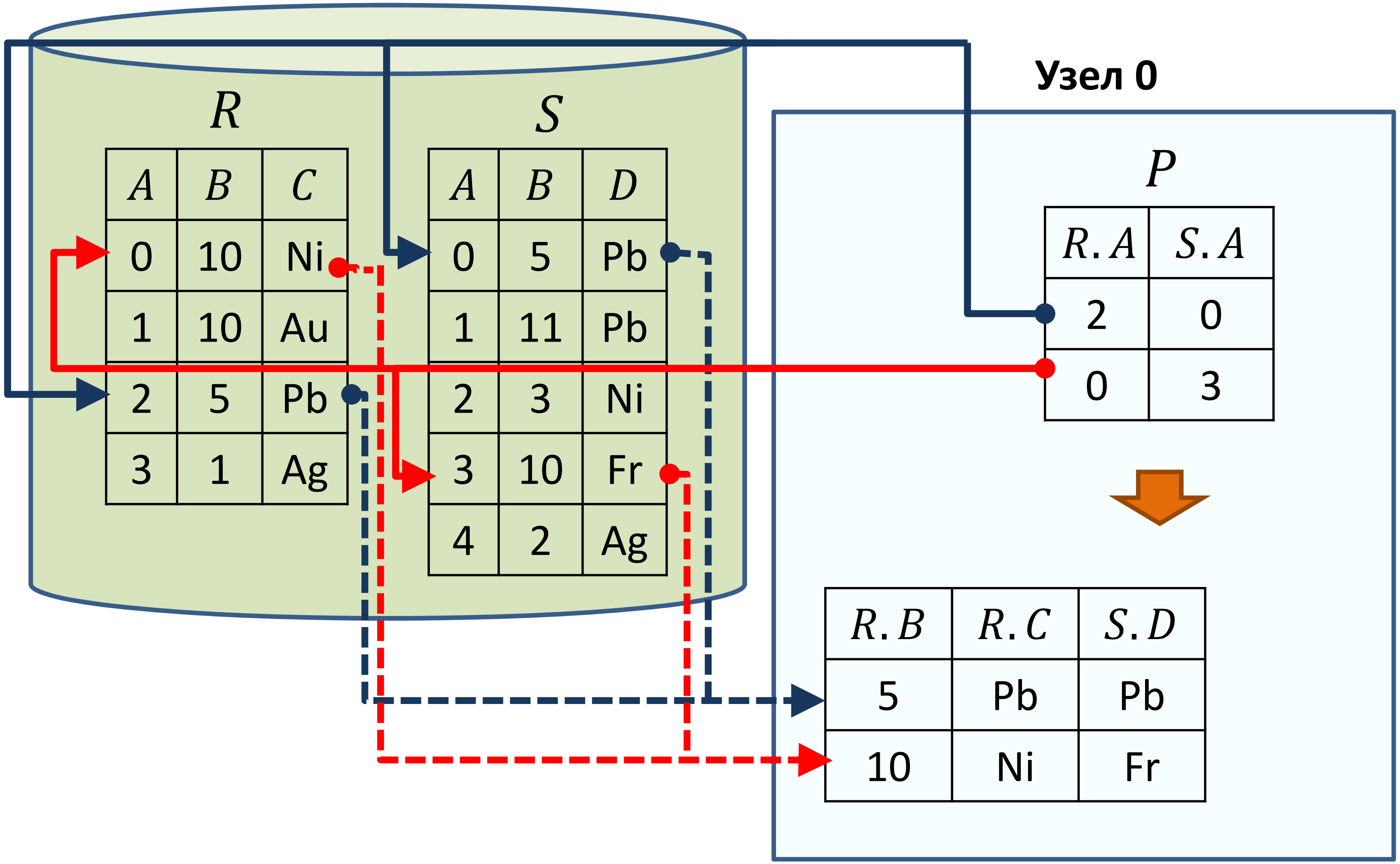
P^1

$R.A$	$S.A$
0	3

P

$R.A$	$S.A$
2	0
0	3

Выполнение операции естественного соединения



Хэш-индекс

- Хеш-индекс позволяет использовать один колоночный индекс для индексирования нескольких атрибутов одного отношения.
- Пусть задано отношение $R(A^*, B_1, \dots, B_u, C, \dots)$. Пусть задана хеш-функция $h : \mathfrak{D}_{B_1} \times \dots \times \mathfrak{D}_{B_u} \to \mathbb{Z}_{\geq 0}$. *Хеш-индексом* $I_h(A^*, H)$ атрибутов $B_1, \dots, B_u$ отношения R называется упорядоченное отношение, удовлетворяющее тождеству:

$$I_h = \tau_H\left(\pi_{A,h(B_1,\dots,B_u)\to H}(R)\right)$$

- *Фрагментация хеш-индекса* осуществляется на основе доменно-интервального принципа.

Операция естественного соединения с использованием хэш-индекса

- Пусть $R(A^*, B_1, \ldots, B_u, C_1, \ldots, C_v)$ и $S(A^*, B_1, \ldots, B_u, D_1, \ldots, D_w)$.
- Определим Q как результат операции $R \bowtie S$.
- Пусть имеется два хэш-индекса для атрибутов $B_1, \ldots, B_u$, построенные с помощью одной и той же хеш-функции h для которых задана доменно-интервальная фрагментация степени k:

$$I_{R,h} = \bigcup_{i=0}^{k-1} I_{R,h}^i \qquad I_{S,h} = \bigcup_{i=0}^{k-1} I_{S,h}^i$$

- Положим для всех $i = 0, \ldots, k-1$

$$P^i = \pi_{I_{R,h}^i.A \to A_R,\ I_{S,h}^i.A \to A_S} \left(I_{R,h}^i \underset{(I_{R,h}^i.H = I_{S,h}^i.H)}{\bowtie} I_{S,h}^i \right) \qquad P = \bigcup_{i=0}^{k-1} P^i$$

$$Q = \{ \left(\&_R(p.A_R).B_1, \ldots, \&_R(p.A_R).B_u, \&_S(p.A_S).D_1, \ldots, \&_S(p.A_S).D_w \right) \mid$$
$$p \in P, \left(\&_R(p.A_R).B_1, \ldots, \&_R(p.A_R).B_u \right) = \left(\&_S(p.A_S).B_1, \ldots, \&_S(p.A_S).B_u \right) \}$$

Пример операции естественного соединения с хэш-индексом

Домен $\mathfrak{D}_H$

(-∞ [120 +∞)

$\varphi_{\mathfrak{D}_H} = 0$ $\varphi_{\mathfrak{D}_H} = 1$

$$\boldsymbol{h(b_1, b_2) = b_1 + ACSII(b_2)}$$

R

A	B_1	B_2	C
0	10	a	Ni
1	9	b	Au
2	5	s	Pb
3	1	s	Ag

S

A	B_1	B_2	D
0	5	s	Pb
1	11	y	Pb
2	3	d	Ni
3	10	a	Fr
4	2	a	Ag

Узел 0

$I^0_{R,h}$

A	H
0	107
1	107
3	116

$I^0_{S,h}$

A	H
4	99
2	103
3	107

Узел 1

$I^1_{R,h}$

A	H
2	120

$I^1_{S,h}$

A	H
0	120
1	132

Пример операции естественного соединения с хэш-индексом

Узел 0

$I^0_{R,h}$

A	H
0	107
1	107
3	116

$I^0_{S,h}$

A	H
4	99
2	103
3	107

P^0

$R.A$	$S.A$
0	3
1	3

P^1

$R.A$	$S.A$
2	0

Узел 1

$I^1_{R,h}$

A	H
2	120

$I^1_{S,h}$

A	H
0	120
1	132

P^1

$R.A$	$S.A$
2	0

Пример операции естественного соединения с хэш-индексом

Узел 0

P^0

$R.A$	$S.A$
0	3
1	3

P^1

$R.A$	$S.A$
2	0

P

$R.A$	$S.A$
0	3
1	3
2	0

Пример операции естественного соединения с хэш-индексом

R

A	B_1	B_2	C
0	10	a	Ni
1	9	b	Au
2	5	s	Pb
3	1	s	Ag

S

A	B_1	B_2	D
0	5	s	Pb
1	11	y	Pb
2	3	d	Ni
3	10	a	Fr
4	2	a	Ag

Узел 0

P

$R.A$	$S.A$
0	3
~~1~~	~~3~~
2	0

$R.B_1$	$R.B_2$	$R.C$	$S.D$
10	a	Ni	Fr
5	s	Pb	Pb

Преимущества распределенных колоночных индексов

- Позволяют исключить обмены при выполнении реляционных операций
- Допускают эффективную балансировку загрузки
- Эффективно утилизируют многоядерные ускорители:
  1. Разжать сегмент
  2. Выполнить операцию
  3. Сжать сегмент

Слабые стороны распределенных колоночных индексов

- Применимы только к отдельным реляционным операциям
- Нельзя создать колоночные индексы на все случаи жизни (для одной базы данных их существует бесконечно много)

Выводы

- Колоночные индексы создаются администратором под определенное приложение баз данных для часто повторяющихся ресурсоемких операций
- Если в запросе встречается операция, для которой отсутствуют необходимые колоночные индексы, то она выполняется обычным образом (без использования суперкомпьютера)

Спасибо за внимание!