

## Алгоритм выбора архитектуры параллельной системы баз данных по критерию стоимости

77-30569/270497

# 12, декабрь 2011

Григорьев Ю. А., Плужников В. Л.

УДК 004.657

МГТУ им. Н.Э. Баумана

[grigorev@iu5.bmstu.ru](mailto:grigorev@iu5.bmstu.ru)

[VPluzhnikov@croc.ru](mailto:VPluzhnikov@croc.ru)

### Оценка стоимости параллельных систем базы данных

Для тех проектов построения информационных систем, для которых важен экономический эффект, должна выбираться архитектура системы с минимальной совокупной стоимостью владения.

Совокупная Стоимость Владения (ТСО - total cost of ownership) - это методика расчета, созданная чтобы помочь потребителям и руководителям предприятий определить прямые и косвенные затраты и выгоды, связанные с любым компонентом компьютерных систем.[11] Также основной целью подсчета стоимости владения, кроме выявления избыточных статей расхода, является оценка возможности возврата вложенных в информационные технологии средств.

В процессе анализа оценки совокупной стоимости владения (ССВ) архитектуры параллельных систем баз данных должны учитываться соответствующие затраты. В работе предлагается проводить оценку ССВ архитектуры на основании следующей формулы:

$$C_{ССВ} = C_O + C_{ПО} + \sum_T C_{Эксплуатации} + C_{Сервис.О} + C_{Сервис.ПО} + C_{Обслуживания}, \quad (1)$$

где

$C_O$  - изначальная стоимость оборудования, которая включает в себя стоимость серверов, систем хранения данных и активной сетевой инфраструктуры, обеспечивающей обмен данными в комплексе,

$C_{ПО}$  - изначальная стоимость программного обеспечения комплекса,

$C_{Эксплуатации}$  - стоимость эксплуатации комплекса в год,

$C_{Сервис.О}$  - стоимость сервисной поддержки оборудования у производителя в год,

$C_{Сервис.ПО}$  - стоимость сервисной поддержки программного обеспечения комплекса у производителя в год,

$C_{Обслуживания}$  - стоимость содержания персонала по обслуживанию комплекса в год и стоимость не прямых затрат на обслуживание комплекса,  
 $T$  - предполагаемое время эксплуатации комплекса в годах.

Детализируем значение стоимости эксплуатации комплекса

$$C_{Эксплуатации} = C_{Эл} + C_{Конд} + C_{Эксп.Др.}, \quad (2)$$

где

$C_{Эл}$  - стоимость электроснабжения комплекса в год,

$C_{Конд}$  - стоимость кондиционирования серверной комнаты в год,

$C_{Эксп.Др.}$  - стоимость обслуживания других инженерных систем в год.

Детализируем значение стоимости обслуживания комплекса

$$C_{Обслуживания} = C_{Персонала} + C_{Аварии} + C_{Обсл.Др.} \quad (3)$$

где

$C_{Персонала}$  - стоимость содержания обслуживающего персонал в год,

$C_{Аварии}$  - стоимость устранения аварийных ситуаций на комплексе в год,

$C_{Обсл.Др.}$  - стоимость дополнительных расходов на обслуживание.

На основе формул (2) и (3), можно переписать формулу совокупной стоимости (1) как

$$C_{ССВ} = C_0 + C_{ПО} + \sum_T \left[ (C_{Эл} + C_{Конд} + C_{Др.}) + C_{Сервис.О} + C_{Сервис.ПО} + (C_{Персонала} + C_{Аварии} + C_{Обсл.Др.}) \right] \quad (4)$$

В расчетах стоимости оборудования необходимо учитывать специфику увеличения стоимости оборудования комплекса в зависимости от изменений технических характеристик.

Ниже приводятся формулы для оценки стоимости ПСБД, состоящей из нескольких SMP-систем (рис. 1). На рис. 1 введены следующие обозначения:

$nAMP$  – число процессоров в одной SMP-системе,

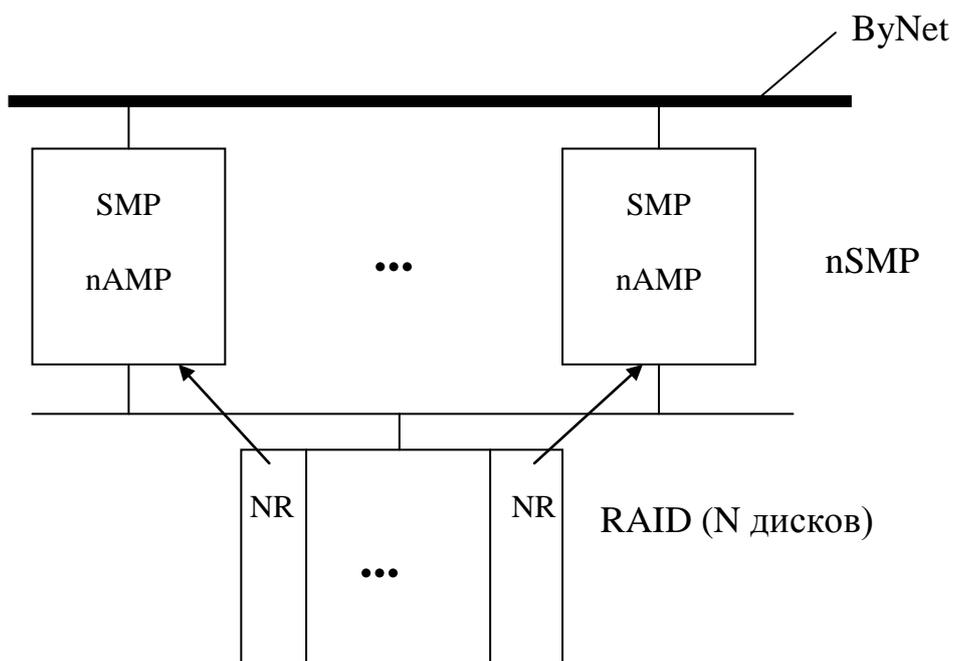
$nSMP$  – число SMP-систем (вычисляется автоматически  $nSMP = n / nAMP$ ),

$NR$  – число дисков, закреплённых за одной SMP-системой. Считается, что шина ввода/вывода очень быстродействующая.

$$NR = N / nSMP \quad (5)$$

Такая конфигурация позволяет исследовать следующие архитектуры:

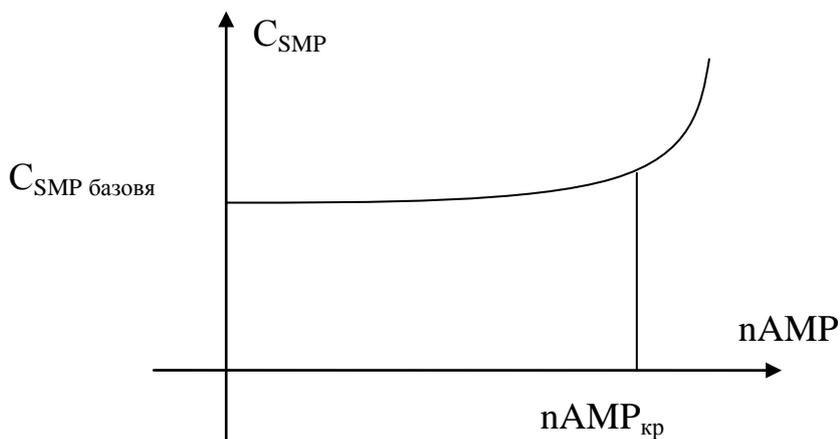
- SE (одна SMP-система),  $n = nAMP$  (т.е.  $nSMP = 1$ ),
- CE (кластер SMP-систем),  $nAMP > 1$  и  $n > nAMP$  (т.е.  $nSMP > 1$ ),
- SN (MPP-система с одним процессором в узле),  $nAMP = 1$  и  $n > nAMP$  (т.е.  $nSMP > 1$ ).



**Рис. 1. Общая схема комплекса, состоящего из нескольких SMP систем.**

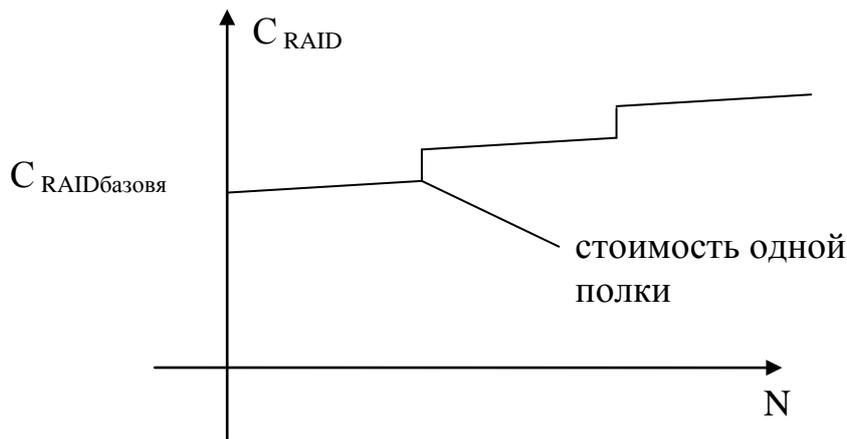
Формулы для оценки стоимости ПСБД определяются особенностями зависимости стоимости системы от числа процессоров и числа дисков в дисковом массиве.

На рис. 2 показана зависимость стоимости SMP-системы от числа процессоров.



**Рис. 2 Зависимость стоимости SMP-системы от числа процессоров.**

На рис. 3 показана зависимость стоимости RAID-массива от числа дисков [12].



**Рис. 3. Зависимость стоимости RAID-массива от числа дисков.**

Допуская, что комплексы с различными архитектурами требуют одинаковые инженерные системы и одинаковую численность обслуживающего персонала, а также то, что системы имеют соизмеримые коэффициенты готовности систем, можно переписать формулу стоимости совокупного владения комплексом, оптимизировав ее для сравнения архитектур параллельных систем баз данных, следующим образом:

$$C_{ССВ.ПСУБД} = C_O + C_{ПО} + \sum_T [(C_{Эл} + C_{Конд}) + C_{Сервис.ПО}] \quad (6)$$

Для сравнительной оценки стоимости различных архитектур параллельных систем баз данных предлагается использовать оценку затрат ежемесячного ССВ комплекса на протяжении пяти лет без модернизации комплекса с выделением следующих компонентов ПСБД: SMP-узлов, системы хранения и коммутационной сети. Таким образом, оценку стоимости определяется по формуле:

$$C_{ССВ.ПСУБД.М} = \frac{(C_{СХД}(N) + nSMP \times C_{SMP}(nAMP) + C_{SW}(nSMP) + C_{О.ДР}) + N_{CPU} \times C_{ПО} + (C_{СХД.Эл}(N) + nSMP \times C_{SMP.Эл}(nAMP) + C_{SW.Эл}(nSMP) + C_{Эл.Др}) + (C_{СХД.Конд}(N) + nSMP \times C_{SMP.Конд}(nAMP) + C_{SW.Конд}(nSMP) + C_{Конд.Др}) + K_{CPU} \times C_{Сервис.ПО}}{60}, \quad (7)$$

где

$C_{СХД}(N)$  – стоимость системы хранения данных, зависящая от числа дисков и дисковых полок в системе хранения ,

$C_{SMP}(nAMP)$  – стоимость SMP-сервера с количеством AMP-процессоров nAMP,

$C_{SW}(nSMP)$  – стоимость коммутатора сети ByNet на nSMP узлов в системе,

$C_{O.Др}$  – стоимость дополнительного оборудования в комплексе (сеть хранения данных, терминальные системы и т.п.),  
 $C_{СХД.Эл}(N)$  – стоимость электроснабжения системы хранения данных в год,  
 $C_{SMP.Эл}(nAMP)$  – стоимость электроснабжения SMP-сервера с количеством AMP-процессоров  $nAMP$ ,  
 $C_{SW.Эл}(nSMP)$  – стоимость электроснабжения коммутатора сети ByNet на  $nSMP$  узлов в системе,  
 $C_{Эл.Др}$  – стоимость электроснабжения дополнительного оборудования в комплексе,  
 $C_{СХД.Конд}(N)$  – стоимость теплоотвода от системы хранения данных в год,  
 $C_{SMP.Конд}(nAMP)$  – стоимость теплоотвода SMP-сервера с количеством AMP-процессоров  $nAMP$ ,  
 $C_{SW.Конд}(nSMP)$  – стоимость теплоотвода коммутатора сети ByNet на  $nSMP$  узлов в системе,  
 $C_{Конд.Др}$  – стоимость теплоотвода дополнительного оборудования в комплексе.

Оценка ССВ в пересчете затрат на месяц позволяет рассчитать рентабельность системы и упрощает оценку построения систем по лизинговой схеме.

### Алгоритм выбора архитектуры ПСБД

Учитывая специфику сравнения архитектур ПСБД и особенности стоимостной оценки, предлагается использовать следующий алгоритм для выбора архитектуры ПСБД.

#### Шаг 1. Рассчитать число дисков в RAID-массиве.

Расчёт числа дисков проводится по формуле (8)

$$N = \left\lceil \frac{Q}{Q_D \cdot P_D} \times k_{RAID} \right\rceil + 2 \cdot k_{enc}, \quad (8)$$

где

$Q$  – общий объём хранимых данных (фактов и измерений),

$Q_D$  – объём диска,

$P_D$  – доля заполнения диска.

$k_{RAID}$  – коэффициент, учитывающий использование технологии RAID для защиты данных от физического отказа дисков. Значения данного коэффициента приведены в таблице 1,

$2 \cdot k_{enc}$  – коэффициент, учитывающий использование технологии горячего резервирования дисков (hot spare).

Значения коэффициента  $k_{RAID}$ 

Тип RAID	Значения $k_{RAID}$
RAID 0+1, RAID 1+0	2
RAID 5 (3+1)	$\left[ \frac{Q}{3 \cdot Q_D \cdot p_D} \right]$
RAID 5 (7+1)	$\left[ \frac{Q}{7 \cdot Q_D \cdot p_D} \right]$
RAID 6 (6+2)	$\left[ \frac{Q}{3 \cdot Q_D \cdot p_D} \right]$
RAID 6 (14+2)	$\left[ \frac{Q}{7 \cdot Q_D \cdot p_D} \right]$

**Шаг 2. Оценить стоимость дискового массива  $C_{СХД}(N)$ .**

На данном шаге проводится запрос стоимости конфигурации системы хранения данных у официальных дистрибуторов оборудования.

**Шаг 3. Проанализировать запросы к хранилищу данных.**

Для каждого  $i$ -го запроса

- 1) определить количество измерений, по которым выполняется поиск ( $K_i$ ),
- 2) оценить число записей таблиц измерений в запросе ( $VP_{ij}, j = 1 \dots K_i$ ),

- 3) рассчитать среднее значение  $VP_i = \sqrt[K_i]{\prod_{j=1}^{K_i} VP_{ij}}$ .

Эти данные занести в табл. 2 и назначить граничные значения для среднего времени выполнения этих запросов.

Таблица 2

**Сводная таблица параметров запросов с граничными значениями для среднего времени их выполнения**

№ запроса	VP	K	Граничное значение для среднего времени выполнения
1	$VP_1$	$K_1$	$T_1$
2	$VP_2$	$K_2$	$T_2$
...			
U	$VP_U$	$K_U$	$T_U$

**Шаг 4. Положить  $n=1$  и  $nAMP=1$** 

Это соответствует самой дешёвой конфигурации (одна SMP-система с одним процессором).

### **Шаг 5. Рассчитать среднее время выполнения запросов.**

Рассчитать среднее время ( $M$ ) для всех запросов из табл. 2, используя соответствующие формулы [10]. Если для какого-либо запроса время его выполнения превышает граничное значение, то перейти к **Шагу 6**, иначе перейти к **Шагу 8**.

### **Шаг 6. Проверить $nAMP$ .**

Если для текущего значения  $nAMP$  перегружается диск массива RAID (дальнейшее увеличение  $nAMP$  не приведёт к уменьшению времени выполнения запросов) или  $nAMP > nAMP_{кр}$  (см. рис. 2), то перейти к **Шагу 7**, иначе положить  $n := (n/nAMP) \times (nAMP + 1)$  – сохраняем число SMP-систем,  $nAMP = nAMP + 1$  – увеличиваем число процессоров в каждой SMP-системе, перейти к **Шагу 5**.

### **Шаг 7. Увеличить число SMP-систем.**

Положить  $n := (n/nAMP) + 1, nAMP = 1$ , перейти к **Шагу 5**.

**Шаг 8. Полученная конфигурация является оптимальной ( $nAMP, nSMP = (n/nAMP)$ ).** Полученные значения  $n, nAMP, nSMP$  необходимо использовать для расчета оценки ССВ архитектуры ПСБД по формуле (7). Завершить алгоритм.

Приведённый выше алгоритм путём последовательного наращивания  $nAMP$  и  $nSMP$  упорядочивает параллельные системы баз данных (ПСБД) с архитектурами SE, CE и SN по возрастанию их стоимости. Это следует из рисунка 2 и описания шага 6.

В этой последовательности можно выделить подпоследовательность, в которой ПСБД упорядочены по убыванию среднего времени выполнения запросов. Действительно, для ПСБД с параметрами  $nAMP$  и  $nSMP$  существует конфигурация ПСБД с параметрами  $nAMP$  и  $nSMP + 1$  (если до этого не было найдено оптимальное решение). Но в силу выражений для среднего времени выполнения запросов [10] эта конфигурация строго лучше по времени, чем предыдущий вариант, т.к. в алгоритме не исследуются системы с перегруженным разделяемым ресурсом (см. шаг 6). Этот вывод также следует из свойства внешней монотонности систем массового обслуживания [13].

Так как в алгоритме последовательно анализируются все варианты ПСБД, то за конечное число итераций будет найдено оптимальное решение.

## ЛИТЕРАТУРА

1. М. Тамер Оззу, Патрик Валдуриз. Распределенные и параллельные системы баз данных: [Электронный ресурс]. [http://citforum.ru/database/classics/distr\_and\_paral\_sdb/]. Проверено 26.11.2010.
2. Соколинский Л. Б., Цымблер М. Л. Лекции по курсу "Параллельные системы баз данных": [Электронный ресурс]. [http://pdbc.susu.ru/CourseManual.html]. Проверено 04.12.2010.
3. Григорьев Ю.А., Плужников В.Л. Оценка времени соединения таблиц в параллельной системе баз данных// Информатика и системы управления. – 2011. - № 1. – С. 3-16.

4. Лисянский К., Слободяников Д. СУБД Teradata® для ОС UNIX®: [Электронный ресурс]. [<http://citforum.ru/database/kbd98/glava5.shtml>]. Проверено 14.03.2011.
5. Кузнецов С. Essential Modelling Options: [Электронный ресурс]. [[http://citforum.ru/database/digest/dig\\_1612.shtml](http://citforum.ru/database/digest/dig_1612.shtml)]. Проверено 14.03.2011.
6. Лев Левин. Teradata совершенствует хранилища данных: [Электронный ресурс]. [<http://www.pcweek.ru/themes/detail.php?ID=71626>]. Проверено 26.11.2010.
7. Григорьев Ю.А., Плутенко А.Д. Теоретические основы анализа процессов доступа к распределённым базам данных. - Новосибирск: Наука, 2002. – 180 с.
8. Миллер Р., Боксер Л. Последовательные и параллельные алгоритмы. Общий подход. – М.: БИНОМ. Лаборатория знаний, 2006. – 406 с.
9. Григорьев Ю.А., Плужников В.Л. Оценка времени соединения таблиц в параллельной системе баз данных // Информатика и системы управления. – 2011. - № 1. – С. 3-16.
10. Григорьев Ю.А., Плужников В.Л. Анализ времени обработки запросов к хранилищу данных в параллельной системе баз данных // Информатика и системы управления. – 2011. - № 2. – С. 94-106.
11. П. Тарасенко Расчет и распределение затрат [Электронный ресурс]. [<http://www.eg-online.ru/article/52214/>] Проверено 23.10.2011.
12. John P. Desmond Infrastructure: storage resource management software and SAN architecture seen lowering TCO // Journal Software Magazine Volume 22 Issue 2, 2002 P. 19-20.
13. Штоян Д. Качественные свойства и оценки стохастических моделей. – М.: Мир, 1979. – 268 с.

## Cost-criterion selection algorithm of parallel system database architecture

77-30569/270497

# 12, December 2011

Григорьев Ю.А., Pluzhnikov V.L.

Bauman Moscow State Technical University

[grigorev@iu5.bmstu.ru](mailto:grigorev@iu5.bmstu.ru)

[VPluzhnikov@croc.ru](mailto:VPluzhnikov@croc.ru)

In the article the authors suggest expressions which make it possible to estimate the aggregate value of possessing database parallel systems (DPS). The authors developed an algorithm of DPS selection which allows for specific character of the comparison of DPS architectures and special feature of cost estimates. By means of sequential growth of the number of AMP-processors and SMP-units the algorithm orders the database parallel systems with SE, CE and SN architectures according to their cost increase. It is shown that in this sequence it is possible to isolate the subsequence in which DPS are ordered according to descending ordering of the mean time of responding to requests. As in the algorithm all DPS versions are consecutively analyzed, in the finite number of iterations optimal solution will be found

---

**Publications with keywords:** [query processing time](#), [Laplace-Stieltjes transformation in concurrent data base system](#), [architecture of concurrent data base system](#)

**Publications with words:** [query processing time](#), [Laplace-Stieltjes transformation in concurrent data base system](#), [architecture of concurrent data base system](#)

---

### Reference

1. M. Tamer Ozzu, Patrik Valduriz, <[http://citforum.ru/database/classics/distr\\_and\\_paral\\_sdb/](http://citforum.ru/database/classics/distr_and_paral_sdb/)>.
2. Sokolinskii L. B., Tsymler M. L., <<http://pdbc.susu.ru/CourseManual.html>>.
3. Grigor'ev Yu.A., Pluzhnikov V.L., Informatika i sistemy upravleniia 1 (2011) 3-16.
4. Lisianskii K., Slobodianikov D., DBMS Teradata® for OS UNIX®, <<http://citforum.ru/database/kbd98/glava5.shtml>>.
5. Kuznetsov S., Essential Modelling Options, <[http://citforum.ru/database/digest/dig\\_1612.shtml](http://citforum.ru/database/digest/dig_1612.shtml)>.
6. Lev Levin, <<http://www.pcweek.ru/themes/detail.php?ID=71626>>.

7. Grigor'ev Iu.A., Plutenko A.D., Theoretical fundamentals of the analysis of the processes of access to distributed databases, Novosibirsk, Nauka, 2002, 180 p.
8. Miller R., Bokser L., The serial and parallel algorithms. The general approach, Moscow, BINOM. Laboratoriia znanii, 2006, 406 p.
9. Grigor'ev Iu.A., Pluzhnikov V.L., Informatika i sistemy upravleniia 1 (2011) 3-16.
10. Grigor'ev Iu.A., Pluzhnikov V.L., Informatika i sistemy upravleniia 2 (2011) 94-106.
11. P. Tarasenko, <<http://www.eg-online.ru/article/52214/>>.
12. John P., Desmond Infrastructure: storage resource management software and SAN architecture seen lowering TCO, Journal Software Magazine 22 (2) (2002) 19-20.
13. Shtoiian D., Qualitative properties and evaluation of stochastic models, Moscow, Mir, 1979, 268 p.