

ЭЛЕМЕНТЫ РАСПРЕДЕЛЁННОЙ СИСТЕМЫ ИМЁН РЕЙЗЕРА И ИХ ВЗАИМОДЕЙСТВИЕ

Кулаков Ю.А., Чубин И.М

kulakov@comsys.ntu-kpi.kiev.ua, imchubin@mail.ru

Национальный Технический Университет Украины "КПИ"

Kulakov Y., Chubin I. ELEMENTS OF THE DISTRIBUTED REISER NAMING SYSTEM AND THEIR INTERACTION. The work contains the notions of distributed Reiser namespace and distributed Reiser naming system; defines elements of distributed Reiser naming system, such as name resolver, name agent and name server. The scheme of Reiser naming system elements is described. Distributed Reiser naming system is more scalable than centralized one.

Введение

Одной из основных задач, определяющих эффективность использования распределённых систем, является задача поиска.

В работах Ганса Рейзера (см. например, [1]) предлагается расширить традиционное иерархическое пространство имён, которое строится только с помощью примитива порядка, примитивом группировки. Это позволит повысить эффективность поиска информации в файловой системе и даст возможность использовать новые подходы к поиску информации. Предлагаемая Рейзером система имён реализованна в виде файловой системы и является централизованной.

В работе предлагается структура распределённой системы имён Рейзера, которая отличается от централизованной системы имён высокой масштабируемостью.

Распределённое пространство имён Рейзера

Пространство имён Рейзера может быть как централизованным, так и распределённым. В централизованном пространстве имён вся информация, описывающая элементы пространства имён, а также связей между ними находится в пределах одного компьютера. В том случае, если пространство имён Рейзера распределёно, информация этого пространства имён хранится не на одном компьютере, а на множестве.

Другими словами, можно составить покрытие $\{RZ_i\}$ графа RNS такое, что

$$RNS = \bigcup_{i \in I} RZ_i.$$

В том случае, если ни одна пара графов из множества $\{RZ_i\}$ не пересекается

$$\forall i, j RZ_i \cap RZ_j = \emptyset,$$

покрытие превращается в разбиение.

Области распределённого пространства имён, соответствующие графам RZ_i разбиения графа RNS , назовём *зонами*. Зона Рейзера – часть распределённого пространства имён Рейзера, целиком обслуживаемая одним элементом распределённой системы (сервером имён).

Для того чтобы описать границу зоны пространства имён, достаточно перечислить вершины, которые её *образуют*. Подграф, соответствующий зоне, состоит из вершин, образующих зону, соединяющих их рёбер, а также рёбер, начало которых находится внутри зоны. Другими словами, ребро принадлежит подграфу её зоны в том случае, если его начало является одной из образующих вершин зоны. В том случае если обе вершины ребра относятся к числу образующих зону, ребро называется *внутризональным*; в противном случае ребро называется *межзональным*.

Пример распределённого пространства имён Рейзера, разбитого на несколько зон, приведен на рисунке 1. На рисунке изображена часть графа RNS пространства имён и условно отмечены границы разбиения на зоны.

$$V = \{r, a, b, c, d, e, f, g, h, i, j, k, l, n\}$$

В каждой зоне выделена одна вершина, название которой определяет имя зоны. Пунктиром отмечены границы зон. Например, зону вершины RZ_c образует три вершины: c , g и h . Соответствующий ей подграф выглядит так:

$$RZ_c = \langle V_c; E_c \rangle$$

$$V_c = \{c, g, h, f, j\}$$

$$E_c = \{(c, f), (c, g), (c, h), (g, j)\}$$

Рёбра (c, f) и (g, j) являются межзональными, а (c, g) и (c, h) внутризональными. Вершины f и j входят в граф RZ_c , хотя и не относятся к числу вершин, образующих зону.

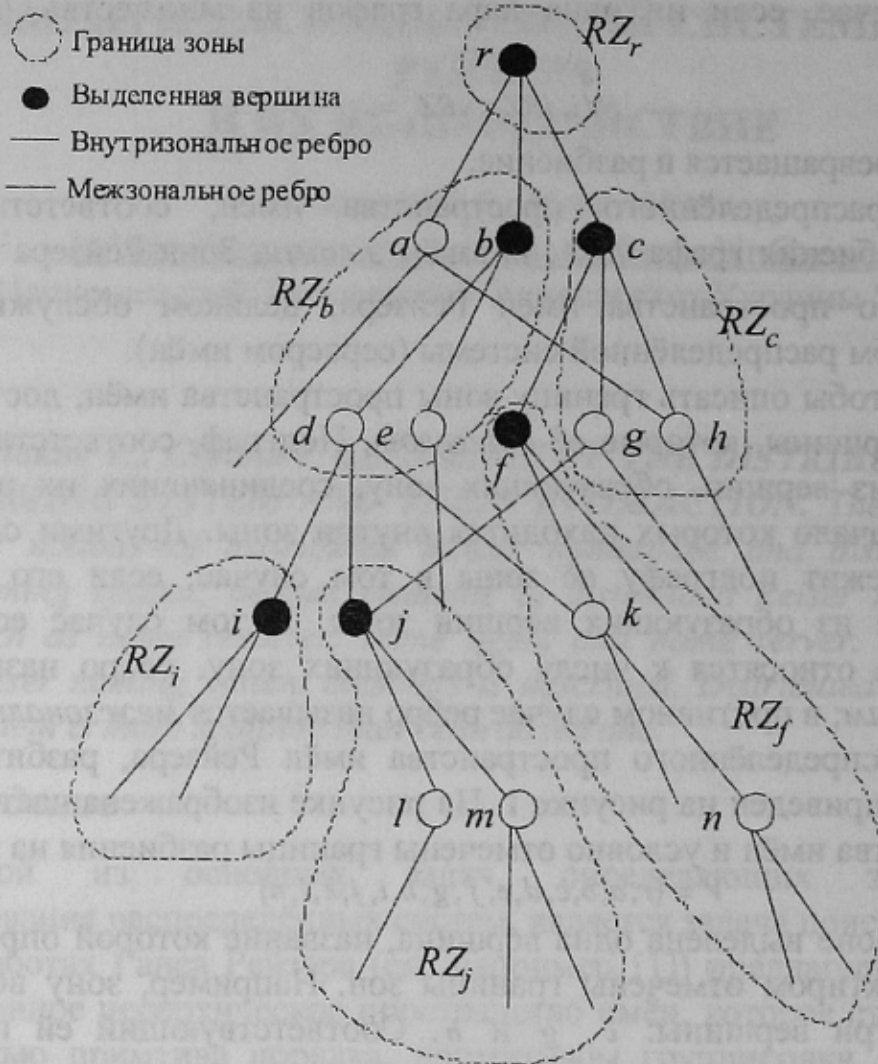


Рисунок 1. Распределённое пространство имён Рейзера

Элементы распределённой системы имён Рейзера и их взаимодействие

В работе распределённого пространства имён Рейзера участвуют элементы, которые можно разделить на три различные группы:

1. Сервера имён.
2. Агенты разрешения имён.
3. Клиенты системы имён.

Сервер имён (авторитативный сервер имён) хранит информацию об одной или нескольких зонах распределённого пространства имён и отвечает на запросы, относящиеся к информации в его зонах.

Агент разрешения имён (кэширующий сервер имён) выполняет обработку запроса на преобразование имени. Он выполняет разрешение имени в соответствии с рассматриваемым ниже алгоритмом разрешения имён Рейзера, опрашивая для этого необходимые сервера имён. Информация, поступающая к агенту разрешения в ходе разрешения имени,

запоминается и хранится им в течение некоторого времени, определяемого временем жизни этой информации, т.е. кэшируется.

Клиент системы имён представляет собой программный модуль, выполняющий функции посредника между прикладным программным обеспечением и распределённой системой имён. Он принимает запрос на разрешение имени и направляет его ближайшему агенту разрешения. После того как разрешение имени выполнено агентом, клиент принимает результат и возвращает его источнику запроса.

Процесс взаимодействия элементов системы имён выглядит так.

Участники:

- прикладной процесс (p);
- клиент системы имён (c);
- агент разрешения имён (a);
- множество серверов имён ($S = \{s_i\}$), обслуживающих зоны RZ_i распределённого пространства имён RNS .

1. $p \rightarrow c$:

Прикладной процесс передаёт клиенту запрос на разрешение.

2. $c \rightarrow a$:

Клиент передаёт агенту запрос на разрешение.

3. $a \rightarrow \{S\}; \{s\} \rightarrow a$:

Агент опрашивает сервера в соответствии с алгоритмом разрешения имён Рейзера.

4. $a \rightarrow c$:

Агент возвращает результат разрешения имени клиенту;

5. $c \rightarrow p$:

Клиент возвращает результат разрешения имени прикладному процессу.

Основную работу по разрешению имени выполняет агент. Он разбивает имя N на элементы и последовательно обрабатывает их, начиная с первого. Информация, полученная в результате обработки очередного имени, используется для обработки последующего. Результат обработки последнего элемента имени является результатом разрешения путевого имени I . Он передается агентом разрешения имён a клиенту c , который в свою очередь возвращает его прикладному процессу.

Рассмотрим пример взаимодействия элементов системы имён при разрешении имени. Пусть распределённое пространство имён содержит имена $|a|$, $|b|$ и $|x|$, которые разбиты по зона следующим образом:

$$a \in RZ_a; b \in RZ_b; x \in RZ_x.$$

Зоны обслуживаются серверами

$$RZ_a \rightarrow s_a; RZ_b \rightarrow s_b; RZ_x \rightarrow s_x.$$

Необходимо разрешить имя $|[a b]/x|$.

Взаимодействие элементов распределённой системы при разрешении имени можно схематически изобразить как на рисунке 2.

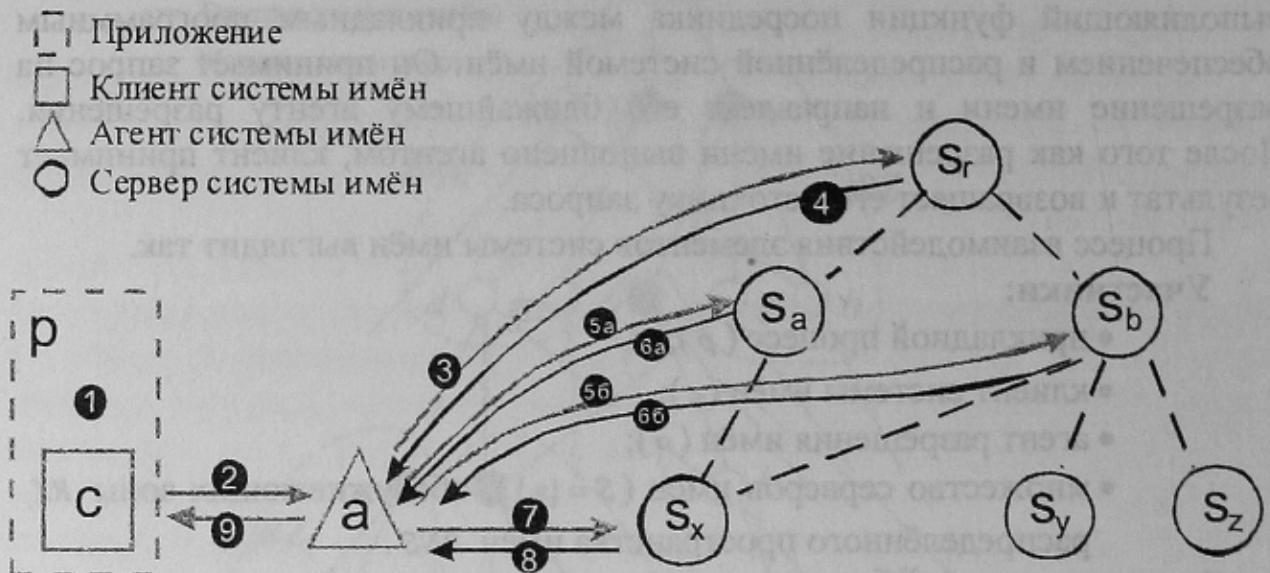


Рисунок 2. Взаимодействие элементов распределённой системы имён Рейзера при разрешении имени

1. Путевое имя $|/[a\ b]/c|$ передается прикладным процессом p клиенту системы имён для преобразования.
2. Клиент c передает имя агенту разрешения имён a .
3. Клиент пытается разрешить имя основываясь на данных из кэша, и если это невозможно, обращается для разрешения имени к серверам имён. Если никакой информации по данному имени у агента нет, разрешение начинается с корневого сервера.

Агент направляет запрос корневому серверу s_r .

1. Сервер не может полностью разрешить имя и возвращает клиенту ссылки на сервера s_a и s_b .
2. Агент направляет запросы серверам s_a и s_b .
3. Ни один из серверов не может разрешить имя полностью. Каждый из них возвращает частичный ответ. В результате пересечения ответов, агент определяет адрес сервера s_x .
4. Агент направляет запрос серверу s_x .
5. Сервер s_x возвращает агенту ответ — результат разрешения имени.
6. Агент возвращает результат разрешения имени клиенту.

В общем случае схема выглядит аналогично. В зависимости от имени и разбиения пространства имён могут появляться дополнительные обращения к серверам имён.

Выводы

Предложенная структура системы имён Рейзера позволяет реализовать пространство имён в распределённой системе, состоящей из множества взаимодействующих компьютеров. В отличие от централизованной системы имён, предлагаемая распределённая система имён обладает значительно более высоким уровнем масштабируемости. Это позволяет значительно расширить область применения рассматриваемых систем.

Литература

1. Reiser, H.: "Future Vision", www.namesys.com
2. Neuman, B.: "Scale in Distributed Systems." Readings in Distributed Computing Systems, pp. 463-489, Los Alamitos, CA: IEEE Computer Society Press, 1994.
3. Tanenbaum, A. and Van Steen, M.: *Distributed Systems. Principles and Paradigms*, Upper Saddle River, NJ: Prentice Hall, 2002.
4. Wieringa, R. and De Jonge, W.: "Object Identifiers, Keys, and Surrogates-Object Identifiers Revisited." Theory and Practice of Object Systems, vol. 1, no. 2, pp. 101-114, 1995.
5. Albitz, P. and Liu, C.: *DNS and BIND*, Sebastopol, CA: O'Reilly & Associates, 3rd ed., 1998.

Дата надходження до редакції 12.06.2005 р.