

ФЕДЕРАЛЬНАЯ СЛУЖБА
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ(12) **ЗАЯВКА НА ИЗОБРЕТЕНИЕ**

(21)(22) Заявка: 2012144349/08, 18.03.2011

Приоритет(ы):

(30) Конвенционный приоритет:
18.03.2010 US 61/315,351

(43) Дата публикации заявки: 27.04.2014 Бюл. № 12

(85) Дата начала рассмотрения заявки РСТ на
национальной фазе: 18.10.2012(86) Заявка РСТ:
US 2011/029056 (18.03.2011)(87) Публикация заявки РСТ:
WO 2011/116324 (22.09.2011)

Адрес для переписки:

129090, Москва, ул. Б. Спасская, 25, стр.3, ООО
"Юридическая фирма Городиский и Партнеры"

(71) Заявитель(и):

НУОДиБи ИНК. (US)

(72) Автор(ы):

СТАРКИ Джеймс А. (US)

(54) **СИСТЕМА УПРАВЛЕНИЯ БАЗОЙ ДАННЫХ**

(57) Формула изобретения

1. Система управления базой данных, которая обеспечивает пользователям возможность взаимодействовать с базой данных, содержащей данные и метаданные, причем упомянутая система содержит:

А) множество узлов, причем каждый узел содержит:

i) интерфейс между командами высокоуровневого ввода и вывода на пользовательском уровне и командами ввода и вывода на системном уровне, которые управляют последовательностью операций для взаимодействия с базой данных, при этом, в ответ на некоторые команды системного уровня, объекты-атомы генерируют атомы, причем каждый атом содержит заданный фрагмент данных или метаданных, посредством чего набор всех экземпляров атомов совместно определяет все метаданные и данные в базе данных,

ii) средство управления связью для установления канала связи с каждым другим узлом в системе,

iii) метод, реагирующий на системную команду от интерфейса для запрашивания от выбранного узла копии атома, который релевантен запросу, но не присутствует в этом узле,

iv) метод, реагирующий на запрос атома от другого узла для репликации запрашиваемого атома для переноса на запрашивающий узел, посредством чего только атомы, требуемые для завершения запроса, должны располагаться в любом узле

транзакции в любой заданный момент времени, и

в) метод, реагирующий на изменение в атоме на этом узле для репликации этого атома для переноса на каждый другой узел в системе, в котором этот атом является постоянно хранимым,

В) средство, подсоединенное к упомянутому множеству узлов, для обеспечения каналов связи между всеми узлами, и

С) средство, подсоединенное к упомянутому средству связи, для обеспечения долговременного хранилища для информации в атоме, посредством чего коллекция атомов в долговременном хранилище совместно содержит все данные и метаданные в базе данных.

2. Система управления базой данных по п.1, в которой некоторые из объектов-атомов создают атомы каталога, которые отслеживают все атомы, которые являются постоянно хранимыми на этом узле, и местоположение каждого узла, который имеет копию каждого из этих постоянно хранимых атомов.

3. Система управления базой данных по п.2, в которой одним из атомов каталога является атом главного каталога, который реплицируется в каждом узле, подсоединенном к системе, при этом другим из атомов является одиночный атом диспетчера транзакций для базы данных для отслеживания хода транзакций, инициированных интерфейсом, и при этом главный каталог идентифицирует местоположение атома диспетчера транзакций.

4. Система управления базой данных по п.3, в которой база данных на пользовательском уровне содержит именованные таблицы с записями и полями и работает в соответствии с по меньшей мере одной именованной схемой, и другие из объектов-атомов создают:

- i) атом таблицы для каждой таблицы, которая содержит список полей,
- ii) атом схемы устанавливает соответствие между схемой и именованными таблицами, соответствующими этой схеме, и
- iii) атом базы данных, который устанавливает соответствие между базой данных и именованной схемой.

5. Система управления базой данных по п.4, в которой таблицы базы данных на пользовательском уровне содержат идентифицированные записи, содержащие поля и данные, при этом другие объекты-атомы создают атомы данных, которые содержат заранее заданное количество записей данных и атом состояний записи, который управляет версиями записи, и при этом упомянутый атом состояний записи включает в себя массив, который идентифицирует управляемые атомы данных, назначенные этому атому состояний записи, идентификатор каждой версии записи и, для каждой версии записи, информацию об этой версии и местоположении атома данных, содержащего эту версию.

6. Система управления базой данных по п.5, в которой каждый атом данных дополнительно включает в себя адрес и длину каждой записи в атоме данных и записей данных.

7. Система управления базой данных по п.4, в которой таблицы базы данных на пользовательском уровне содержат идентифицированные блобы, при этом другие объекты-атомы создают атомы блоба, которые содержат блоб и атом состояний записи блоба, который включает в себя список всех ассоциированных атомов блоба и местоположение блоба в атоме блоба.

8. Система управления базой данных по п.7, в которой каждый атом блоба включает в себя адрес и длину каждого блоба и блоб.

9. Система управления базой данных по п.1, в которой каждый узел дополнительно содержит:

i) средство обмена сообщениями для генерирования и приема предопределенных сообщений в ответ на команды системного уровня, и

ii) средство для выбора одного из каналов связи для передачи сообщения.

10. Система управления базой данных по п.9, в которой каждый узел содержит средство для асинхронного тестирования эффективности каждого канала связи.

11. Система управления базой данных по п.10, в которой упомянутое средство тестирования включает в себя средство для генерирования сообщения тестового опроса и приема сообщения подтверждения приема тестового опроса для определения времени тестового опроса для каждого канала связи.

12. Система управления базой данных по п.9, в которой упомянутое средство обмена сообщениями включает в себя средство для сериализации атома в сообщение и десериализации сообщения в атом, при этом каждое сообщение асинхронно посылается по выбранному каналу связи.

13. Система управления базой данных по п.12, в которой сообщения посылаются согласно протоколу с управлением, которое сохраняет последовательности сообщений.

14. Система управления базой данных по п.1, в которой отдельные узлы могут быть активными и подсоединенными к системе и могут быть отсоединенными от системы, и один из атомов является атомом главного каталога, который идентифицирует каждый активный узел в системе и который реплицируется в каждом узле, подсоединенном к системе, причем система содержит средства, позволяющие узлу присоединяться к системе, включающие в себя:

i) средство для установления соединения между присоединяющимся узлом и выбранным одним из активных узлов,

ii) средство в присоединяющемся узле для запрашивания выбранного узла на предмет подсоединения,

iii) средство в выбранном узле для обновления своего атома главного каталога для указания присоединяющегося узла в качестве активного узла,

iv) средство в выбранном узле для переноса на присоединяющийся узел сообщения, включающего в себя идентификатор выбранного узла, и для ширококестельной передачи сообщения, указывающего доступность присоединяющегося узла, всем другим узлам в системе и долговременному хранилищу,

v) средство в упомянутом присоединяющемся узле для ответа на сообщение посредством запрашивания копии атома главного каталога с выбранного узла, и

vi) средство в выбранном узле для переноса своего обновленного атома главного каталога на присоединяющийся узел.

15. Система управления базой данных по п.14, в которой упомянутый присоединяющийся узел включает в себя средство, реагирующее на прием обновленного атома главного каталога для запрашивания копий дополнительных атомов в ответ на запросы от интерфейса.

16. Система управления базой данных по п.14, в которой выбранный узел выполняет ширококестельную передачу доступности обновленного главного каталога всем другим узлам и долговременному хранилищу, и каждый другой узел и долговременное хранилище включают в себя средство для подтверждения приема им обновленного главного каталога выбранному узлу.

17. Система управления базой данных по п.1, в которой интерфейс в одном из узлов выполняет запрос, который требует создания нового атома, при этом некоторые из объектов-атомов создают атомы каталога, которые отслеживают атомы, которые являются постоянно хранимыми на этом узле, причем упомянутый один из упомянутых узлов дополнительно содержит:

i) средство в упомянутом атоме каталога, которое отслеживает новый атом,

реагирующий на запрос интерфейса, чтобы вызывать создание атомом каталога этого нового атома без содержимого и объявлять новый атом в качестве председателя для этого атома,

ii) средство в атоме каталога, которое назначает идентификатор объекта новому атому,

iii) средство в атоме каталога для установления статуса для нового атома, который указывает, что нет избыточной копии нового атома и что новый атом недоступен для других узлов,

iv) средство в атоме каталога для широковещательной передачи всем другим соответствующим атомам каталога в системе существования нового атома,

v) средство для заполнения нового атома и выполнения нового атома, доступным в системе.

18. Система управления базой данных по п.17, в которой долговременное хранилище включает в себя средство, реагирующее на сохранение нового атома для отправки сообщения узлу, причем новый атом обновляет свой статус для указания существования избыточной копии.

19. Система управления базой данных по п.1, в которой интерфейс в одном из узлов выполняет запрос атома, который существует только в других узлах, при этом некоторые из объектов-атомов создают атомы каталога, которые отслеживают атомы, которые являются постоянно хранимыми на этом узле и других узлах, причем упомянутая система включает в себя:

i) средство в запрашивающем атоме каталога, реагирующее на запрос интерфейса, который отслеживает запрашиваемый атом в качестве реакции на запрос интерфейса, для создания этого пустого атома,

ii) средство в запрашивающем атоме каталога для установления выбора наиболее быстро реагирующего узла и запроса копии атома от наиболее быстро реагирующего узла,

iii) средство в выбранном узле для отправки копии запрашиваемого атома в сообщении объекта на запрашивающий узел и для широковещательной передачи доступности атома другим узлам, в которых атом,

iv) средство для оценки достоверности принятого сообщения объекта для заполнения пустого атома и для направления сообщения подтверждения приема объекта на выбранный узел.

20. Система управления базой данных по п.19, в которой упомянутый выбранный узел генерирует список ретрансляций всех узлов, содержащих запрашиваемый атом, на которые выбранный узел будет направлять сообщения с других, относящихся к этому атому, на запрашивающий узел.

21. Система управления базой данных по п.20, в которой один из других узлов, которые принимают широковещательное сообщение доступности объекта, включает в себя средство для отправки сообщения подтверждения приема объекта на выбранный узел, причем упомянутый выбранный узел включает в себя средство для обновления списка ретрансляций в ответ на прием каждого сообщения подтверждения приема объекта от другого узла для завершения ретрансляции любых последующих сообщений на запрашивающий узел.

22. Система управления базой данных по п.1, приспособленная для обработки транзакций для гарантирования непротиворечивости базы данных, при этом система включает в себя первый узел, включающий в себя атом, подвергаемый фиксации, причем по меньшей мере один другой узел включает в себя копию этого атома, и копия этого атома включена в долговременное хранилище, причем упомянутая система включает в себя:

A
6
4
3
4
9
2
0
1
2
1
4
3
4
9
A
R
U

R
U
2
0
1
2
1
4
3
4
9
A

i) средство в первом узле для генерирования сообщения предварительной фиксации с идентификатором транзакции для пересылки в долговременное хранилище,
ii) средство, ассоциированное с долговременным хранилищем, реагирующее на сообщение предварительной фиксации, для фиксации транзакции и после этого - широковещательной передачи сообщения фиксации с идентификатором транзакции на все узлы,

iii) средство в первом и по меньшей мере одном другом узле, реагирующее на прием сообщения фиксации, для обновления идентифицированной транзакции.

23. Система управления базой данных, которая позволяет пользователям взаимодействовать с базой данных, содержащей данные и метаданные, причем упомянутая система содержит

А) по меньшей мере один узел транзакции, который предоставляет пользователям доступ к базе данных, и по меньшей мере один узел архивации, который поддерживает архив всей базы данных, причем каждый узел транзакции включает в себя подсистему запросов к базе данных, которая обеспечивает интерфейс между командами запроса высокоуровневого ввода и вывода на пользовательском уровне и командами ввода и вывода на системном уровне, которые управляют последовательностью операций для взаимодействия с базой данных, при этом в ответ на некоторые команды системного уровня объекты-атомы генерируют атомы, причем каждый атом содержит заданный фрагмент данных или метаданных, посредством чего набор всех экземпляров атомов совместно определяет все метаданные и данные в базе данных,

В) сеть системы базы данных, соединяющую между собой все узлы,

С) средство управления связью в каждом из узлов для установления канала связи с каждым другим узлом в системе,

Д) метод в каждом узле транзакции, реагирующий на системную команду от подсистемы запросов к базе данных, для запрашивания копии атома, который релевантен команде запроса, но не присутствует в этом узле,

Е) метод в каждом узле, реагирующий на запрос атома от другого узла, для репликации запрашиваемого атома для переноса на запрашивающий узел, посредством чего только атомы, необходимые для завершения команды запроса, должны располагаться в любом узле транзакции в любой заданный момент времени, и

Ф) метод в каждом узле транзакции, реагирующий на изменение в атоме на этом узле, для репликации этого изменения на каждый другой узел в системе, который содержит копию этого атома.

24. Система управления базой данных по п.23, в которой некоторые из объектов-атомов создают атомы каталога, которые отслеживают все атомы, которые являются постоянно хранимыми на этом узле, и местоположение каждого узла, который имеет копию каждого из постоянно хранимых атомов.

25. Система управления базой данных по п.24, в которой один из атомов каталога является атомом главного каталога, который реплицируется в каждом узле, подсоединенном к системе, при этом другим из атомов является единичный атом диспетчера транзакций для базы данных для отслеживания хода транзакций, инициированных интерфейсом, при этом главный каталог идентифицирует местоположение атома диспетчера транзакций.

26. Система управления базой данных по п.25, в которой база данных на пользовательском уровне содержит именованные таблицы с записями и полями и работает в соответствии с по меньшей мере одной именованной схемой, и другие из объектов-атомов создают:

i) атом таблицы для каждой таблицы, которая содержит список полей,

ii) атом схемы устанавливает соответствие между схемой и именованными таблицами,

соответствующими этой схеме, и

iii) атом базы данных, который устанавливает соответствие между базой данных и именованной схемой.

27. Система управления базой данных по п.26, в которой таблицы базы данных на пользовательском уровне содержат идентифицированные записи, содержащие поля и данные, причем другие объекты-атомы создают атомы данных, которые содержат predetermined количество записей данных, и атом состояний записи, который управляет версиями записи, при этом упомянутый атом состояний записи включает в себя массив, который идентифицирует управляемые атомы данных, назначенные этому атому состояний записи, идентификатор каждой версии записи и, для каждой версии записи, информацию об этой версии и местоположение атома данных, содержащего эту версию.

28. Система управления базой данных по п.27, в которой каждый атом данных дополнительно включает в себя адрес и длину каждой записи в атоме данных и записей данных.

29. Система управления базой данных по п.26, в которой таблицы базы данных на пользовательском уровне содержат идентифицированные блобы, при этом другие объекты-атомы создают атомы блоба, которые содержат блоб и атом состояний записи блоба, который включает в себя список всех ассоциированных атомов блоба и местоположение блоба в атоме блоба.

30. Система управления базой данных по п.29, в которой каждый атом блоба включает в себя адрес и длину каждого блоба и блоб.

31. Система управления базой данных по п.23, в которой каждый узел дополнительно содержит:

i) средство обмена сообщениями для генерирования и приема predetermined сообщений в ответ на команды системного уровня, и

ii) средство для выбора одного из каналов связи для передачи сообщения.

32. Система управления базой данных по п.31, в которой каждый узел содержит средство для асинхронного тестирования эффективности каждого канала связи.

33. Система управления базой данных по п.32, в которой упомянутое средство тестирования включает в себя средство для генерирования сообщения тестового опроса и приема сообщения подтверждения приема тестового опроса для определения времени тестового опроса для каждого канала связи.

34. Система управления базой данных по п.31, в которой упомянутое средство обмена сообщениями включает в себя средство для сериализации атома в сообщение и десериализации сообщения в атом, при этом каждое сообщение асинхронно посылается по выбранному каналу связи.

35. Система управления базой данных по п.34, в которой сообщения посылаются в соответствии с протоколом с управлением, которое сохраняет последовательности сообщений.

36. Система управления базой данных по п.23, в которой отдельные узлы могут быть активными и подсоединенными к системе и могут быть отсоединенными от системы, и один из атомов является атомом главного каталога, который идентифицирует каждый активный узел в системе и который реплицируется в каждом узле, подсоединенном к системе, причем система содержит средства, позволяющие узлу присоединиться к системе, включающие в себя:

i) средство для установления соединения между присоединяющимся узлом транзакции и выбранным одним из активных узлов,

ii) средство в присоединяющемся узле транзакции для запрашивания выбранного узла на предмет подсоединения к системе,

- iii) средство в выбранном узле для обновления его атома главного каталога для указания присоединяющегося узла транзакции в качестве активного узла,
- iv) средство в выбранном узле для переноса сообщения на присоединяющийся узел транзакции, включающего в себя идентификатор выбранного узла, и для широковещательной передачи сообщения, указывающего доступность присоединяющегося узла транзакции, всем другим узлам в системе,
- v) средство в упомянутом присоединяющемся узле транзакции для ответа на сообщение посредством запрашивания копии атома главного каталога у выбранного узла, и
- vi) средство в выбранном узле для пересылки его обновленного атома главного каталога на присоединяющийся узел транзакции.

37. Система управления базой данных по п.36, в которой упомянутый присоединяющийся узел транзакции включает в себя средство, реагирующее на прием обновленного атома главного каталога для запрашивания копий дополнительных атомов в ответ на системные команды от ее подсистемы запросов к базе данных.

38. Система управления базой данных по п.36, в которой выбранный узел выполняет широковещательную передачу доступности обновленного главного каталога на все другие узлы, и каждый другой узел включает в себя средство для подтверждения им приема обновленного главного каталога на выбранный узел.

39. Система управления базой данных по п.23, в которой подсистема запросов к базе данных в одном из узлов транзакции выполняет запрос, который требует создания нового атома, при этом некоторые из объектов-атомов создают атомы каталога, которые отслеживают атомы, которые являются постоянно хранимыми на этом узле, причем упомянутый узел транзакции дополнительно содержит:

- i) средство в упомянутом атоме каталога для отслеживания нового атома, реагирующего на запрос подсистемы запросов к базе данных, вызывающее создание атомом каталога этого нового атома без содержимого и объявляющее новый атом в качестве председателя для этого атома,
- ii) средство в атоме каталога для назначения идентификатора объекта новому атому,
- iii) средство в атоме каталога для установления статуса для нового атома, который указывает, что нет избыточной копии нового атома и что новый атом недоступен для других узлов,
- iv) средство в атоме каталога для широковещательной передачи, на все другие соответствующие атомы каталога в системе, существования нового атома,
- v) средство для заполнения нового атома и для выполнения нового атома доступным для других узлов в системе.

40. Система управления базой данных по п.39, дополнительно включающая в себя средство в каждом узле архивации, реагирующее на сохранение нового атома, для посылки сообщения на новый атом, обновляющего его статус для указания существования избыточной копии.

41. Система управления базой данных по п.23, в которой подсистема запросов к базе данных в одном узле транзакции запрашивает атом, который существует только в других узлах, причем некоторые из объектов-атомов создают атомы каталога, которые отслеживают атомы, которые являются постоянно хранимыми на этом узле и на других узлах, причем упомянутая система включает в себя:

- i) средство, реагирующее на запрос в запрашивающем атоме каталога, которое будет отслеживать запрашиваемый атом для создания этого пустого атома,
- ii) средство в запрашивающем атоме каталога для установления выбора наиболее быстро реагирующего узла и запрашивания копии атома от наиболее быстро реагирующего узла,

iii) средство в выбранном узле для посылки копии запрашиваемого атома в сообщении объекта на запрашивающий узел и для широковещательной передачи доступности атома другим узлам, в которых атом постоянно находится,

iv) средство для оценки достоверности принятого сообщения объекта, для заполнения пустого атома и для направления сообщения подтверждения приема объекта на выбранный узел.

42. Система управления базой данных по п.41, в которой упомянутый выбранный узел генерирует список ретрансляций всех узлов, содержащих запрашиваемый атом, на который выбранный узел будет направлять сообщения другого, относящегося к этому атому, на запрашивающий узел.

43. Система управления базой данных по п.42, в которой каждый из других узлов, которые принимают широковещательное сообщение доступности объекта, включает в себя средство для посылки сообщения подтверждения приема объекта на выбранный узел, причем упомянутый выбранный узел включает в себя средство для обновления списка ретрансляций в ответ на прием каждого сообщения подтверждения приема объекта от другого узла для завершения ретрансляции любых последующих сообщений на запрашивающий узел.

44. Система управления базой данных по п.23, приспособленная для гарантирования непротиворечивости базы данных во время обработки транзакций, при этом система включает в себя первый узел, приводящий к фиксации транзакции, и по меньшей мере один другой узел, включающий в себя узел архивации, включает в себя копию этого включающего атома, причем упомянутая система включает в себя:

i) средство в первом узле для генерирования сообщения предварительной фиксации с идентификатором транзакции для переноса на узел архивации,

ii) средство, ассоциированное с узлом архивации, реагирующее на сообщение предварительной фиксации, для фиксации транзакции и затем - широковещательной передачи сообщения фиксации с идентификатором транзакции на все узлы, и

iii) средство в первом и по меньшей мере одном другом узле, реагирующее на прием сообщения фиксации для обновления идентифицированной транзакции.

45. Система управления базой данных для логической базы данных, содержащей записи данных, организованные в таблицы, к которым должен осуществляться доступ с многочисленных узлов транзакции, которые обрабатывают транзакции, связанные с логической базой данных, причем упомянутая система содержит:

А) средство для синтаксического разбора базы данных на фрагменты, при этом каждый фрагмент хранит часть метаданных и/или данных, относящихся к логической базе данных, для переноса в системе управления базой данных в качестве сериализованных сообщений и для сохранения в качестве десериализованного сообщения,

В) по меньшей мере один узел архивации, который хранит все фрагменты в десериализованном виде в долгосрочном хранилище, таким образом составляя единое хранилище для всей базы данных, при этом каждый узел транзакции содержит:

i) средство для ответа на запросы от пользователей посредством установления последовательности низкоуровневых команд для идентификации фрагментов, которые релевантны запросу,

ii) средство для ответа на низкоуровневые команды посредством получения только тех копий существующих фрагментов, которые релевантны запросу, обрабатываемому в нем, притом что заданный фрагмент может существовать на некотором другом узле или только на узле архивации, и

iii) средство для репликации любого измененного фрагмента на упомянутый по меньшей мере один узел архивации и каждый узел транзакции, в котором постоянно

хранится копия этого фрагмента, посредством чего изменения выполняются в отношении фрагментов в других узлах на равноправной основе и посредством чего любой узел транзакции содержит только те фрагменты, которые относятся к запросам, выполняемым пользователями, осуществляющими доступ к базе данных посредством этого узла транзакции.

RU 2012144349 A

A 64341210 RU