

Claims

What is claimed is:

1. A client apparatus of a memcached system, comprising:
memory;
5 one or more processors coupled with the memory; and
memcached logic operated by the one or more processors to:
receive a request to Get, Set or Update a value corresponding to a key in
the memcached system,
determine, in response to receipt of the request, whether the key results in a
10 hit in a local cache maintained in memory by the memcached logic, and
service the Get, Set or Update request based at least in part on whether a
result of the determine indicates the key results in a hit in the local cache.
2. The client apparatus of claim 1, wherein the memcached logic is to further
allocate or cause to be allocated storage spaces within the memory for the local cache.
- 15 3. The client apparatus of claim 1, wherein for a Get request with the key
results in a hit, service comprises
access the local cache to retrieve one or more addresses of one or more
locations of the memcached system where the value and an associated status flag
of the value are stored;
20 perform a remote direct memory access operation on a server of the
memcached system to retrieve the value and the associated status flag using the
one or more addresses retrieved from the local cache; and
determine whether the value retrieved through the remote direct memory
access operation is stale, based at least in part on the status flag retrieved through
25 the remote direct memory access operation.
4. The client apparatus of claim 3, wherein for a Get request with the key
results in a hit, service further comprises return the retrieved value, in response to a result
of the stale determination indicating the value is not stale.
5. The client apparatus of claim 3, wherein for a Get request with the key
30 results in a hit, service further comprises:

in response to a result of the stale determination indicating the value is stale,
transmit the Get request to a server of the memcached system to request the
value;

5 receive second one or more addresses of second one or more locations of
the memcached system where the value and an associated status flag of the value
are stored;

perform a remote direct memory access operation on a server of the
memcached system to retrieve the value and the associated status flag using the
second one or more addresses received;

10 update the local cache to store the third and fourth addresses for the key;

and

return the value received.

6. The client apparatus of claim 1, wherein for a Set request with the key not
results in a hit, service comprises:

15 transmit the Get request to a server of the memcached system to request the
value;

receive one or more addresses of one or more locations of the memcached
system where the value, and an associated status flag of the value are stored;

20 perform a remote direct memory access operation on a server of the
memcached system to retrieve the value and the associated status flag using the
one or more addresses received; and

return the value received.

7. The client apparatus of claim 1, wherein for a Set request, service
comprises:

25 in response to the key results in a hit, invalidate a cache entry in the local
cache that yielded the hit; and

send the Set request to a server of the memcached system.

8. The client apparatus of claim 3, wherein perform a remote direct memory
access comprises first determine a cache entry that yields the hit has not expired.

30 9. A server apparatus of a memcached system, comprising:
memory;

one or more processors coupled with the memory;
network interface card equipped to support remote memory direct access; and
memcached logic operated by the one or more processors to:

5 receive a request to Get an value corresponding to a key in the memcached
system,
determine whether the key results in a hit, and
in response to the key results in a hit, return one or more addresses of one
or more locations of the memcached system where the value and a status flag
associated with the value are stored.

10 10. The server apparatus of claim 9, wherein the memcached logic to further
receive a request to Update a second key in the memcached system with a
second corresponding value,
determine whether the second key results in a hit, and
15 in response to the second key results in a hit, update a second status flag
associated with a first item that yields the hit for the second key with a stale status,
add the first item to a collect list, create a second item for the second key, and
assign the second item with the second value.

11. The server of apparatus of claim 9, wherein the memcached logic to further
periodically perform a maintenance operation to recover memory allocated to items having
20 associated status flags with a stale status.

12. The server of apparatus of claim 9, wherein the memcached logic to further
map physical addresses of the memory to an address space of the memcached system.

13. A method for memory caching data, comprising:
receiving, by a client device of a memocached system, a request to Get, Set,
25 or Update a value corresponding to a key in the memcached system;
determining, by the client device, in response to the receive, whether the
key results in a hit in a local cache maintained in memory of the client device, and
servicing, by the client device, the Get, Set, or Update request based at least
in part on whether a result of the determine indicates the key results in a hit in the
30 local cache.

14. The method of claim 13, wherein further comprising the client device allocating storage spaces within the memory for the local cache.

15. The method of claim 13, wherein for a Get request with the key results in a hit, servicing comprises

5 accessing the local cache to retrieve one or more addresses of one or more locations of the memcached system where the value and an associated status flag of the value are stored;

 performing a remote direct memory access operation on a server of the memcached system to retrieve the value and the associated status flag using the one or more addresses retrieved from the local cache; and

10 determining whether the value retrieved through the remote direct memory access operation is stale, based at least in part on the status flag retrieved through the remote direct memory access operation.

16. The method of claim 15, wherein for a Get request with the key results in a hit, servicing further comprises returning the retrieved value, in response to a result of the stale determination indicating the value is not stale.

17. The method of claim 15, wherein for a Get request with the key results in a hit, servicing further comprises:

20 in response to a result of the stale determination indicating the value is stale, transmitting the Get request to a server of the memcached system to request the value;

 receiving second one or more addresses of second one or more locations of the memcached system where the value and an associated status flag of the value are stored;

25 performing a remote direct memory access operation on a server of the memcached system to retrieve the value and the associated status flag using the second one or more addresses received;

 updating the local cache to store second or more addresses for the key; and returning the value received.

30 18. The method of claim 13, wherein for a Set request with the key not results in a hit, service comprises:

transmitting the Get request to a server of the memcached system to request the value;

receiving one or more addresses of one or more locations of the memcached system where the value and an associated status flag of the value is stored;

performing a remote direct memory access operation on a server of the memcached system to retrieve the value and the associated status flag using the one or more addresses received; and

returning the value received.

19. The method of claim 13, wherein for a Set request, servicing comprises: in response to the key results in a hit, invalidating a cache entry in the local cache that yielded the hit; and sending the Set request to a server of the memcached system.

20. The method of claim 15, wherein performing a remote direct memory access operation comprises first determining that a cache entry that yields the hit has not expired.

21. A method for memory caching data, comprising: receiving, by a server of a memcached system, a request to Get an value corresponding to a key, determining, by the server, whether the key results in a hit, and in response to the key results in a hit, returning, by the server, one or more addresses of one or more locations of the memcached system where the value and a status flag associated with the value are stored.

22. The method of claim 21, further comprising: receiving, by the server, a request to Update a second key in the memcached system with a second corresponding value, determine, by the server, whether the second key results in a hit, and in response to the second key results in a hit, the server, updating a second status flag associated with a first item that yields the hit for the second key with a stale status, adding the first item to a collect list, creating a second item for the second key, and assigning the second item with the second value.

23. The method of claim 21, further comprising periodically performing, by the server, a maintenance operation to recover memory allocated to items having associated status flags with a stale status.

5 24. The method of claim 21, further comprising mapping, by the server, physical addresses of memory of the server to an address space of the memcached system.

25. At least one storage medium comprising a plurality of instructions, in response to execution of the instructions by a client apparatus, to cause the client apparatus to perform any one of the methods of claims 13-18.

10 26. At least one storage medium comprising a plurality of instructions, in response to execution of the instructions by a server apparatus, to cause the server apparatus to perform any one of the methods of claims 19-24.

27. A client apparatus of a memcached system, comprising:
means for receiving a request to Get, Set or Update a value corresponding to a key in the memcached system;
15 means for determining, in response to the receive, whether the key results in a hit in a local cache maintained in memory of the client device; and
means for servicing, the Get, Set or Update request based at least in part on whether a result of the determine indicates the key results in a hit in the local cache.

20 28. A server apparatus of a memcached system, comprising:
means for receiving a request to Get an value corresponding to a key;
means for determining whether the key results in a hit, and
means for responding to the key resulting in a hit, including means for returning one or more address of one or more locations of the memcached system where the value and a status flag associated with the value are stored.

25

CLAIMS

What is claimed is:

1. A client apparatus of a memcached system, comprising:
memory;
one or more processors coupled with the memory; and
memcached logic operated by the one or more processors to:
receive a request to Get, Set or Update a value corresponding to a key in the memcached system,
determine, in response to receipt of the request, whether the key results in a hit in a local cache maintained in memory by the memcached logic, and
service the Get, Set or Update request based at least in part on whether a result of the determine indicates the key results in a hit in the local cache.
2. The client apparatus of claim 1, wherein the memcached logic is to further allocate or cause to be allocated storage spaces within the memory for the local cache.
3. The client apparatus of claim 1, wherein for a Get request with the key results in a hit, service comprises
access the local cache to retrieve one or more addresses of one or more locations of the memcached system where the value and an associated status flag of the value are stored;
perform a remote direct memory access operation on a server of the memcached system to retrieve the value and the associated status flag using the one or more addresses retrieved from the local cache; and
determine whether the value retrieved through the remote direct memory access operation is stale, based at least in part on the status flag retrieved through the remote direct memory access operation.
4. The client apparatus of claim 3, wherein for a Get request with the key results in a hit, service further comprises return the retrieved value, in response to a result of the stale determination indicating the value is not stale.

5. The client apparatus of claim 3, wherein for a Get request with the key results in a hit, service further comprises:

in response to a result of the stale determination indicating the value is stale,
transmit the Get request to a server of the memcached system to request the value;
receive second one or more addresses of second one or more locations of the memcached system where the value and an associated status flag of the value are stored;
perform a remote direct memory access operation on a server of the memcached system to retrieve the value and the associated status flag using the second one or more addresses received;
update the local cache to store the third and fourth addresses for the key; and
return the value received.

6. The client apparatus of claim 1, wherein for a Set request with the key not results in a hit, service comprises:

transmit the Get request to a server of the memcached system to request the value;
receive one or more addresses of one or more locations of the memcached system where the value, and an associated status flag of the value are stored;
perform a remote direct memory access operation on a server of the memcached system to retrieve the value and the associated status flag using the one or more addresses received; and
return the value received.

7. The client apparatus of claim 1, wherein for a Set request, service comprises:
in response to the key results in a hit, invalidate a cache entry in the local cache that yielded the hit; and
send the Set request to a server of the memcached system.

8. The client apparatus of claim 3, wherein perform a remote direct memory access comprises first determine a cache entry that yields the hit has not expired.

9. A server apparatus of a memcached system, comprising:
memory;
one or more processors coupled with the memory;

network interface card equipped to support remote memory direct access; and memcached logic operated by the one or more processors to:

receive a request to Get an value corresponding to a key in the memcached system,

determine whether the key results in a hit, and

in response to the key results in a hit, return one or more addresses of one or more locations of the memcached system where the value and a status flag associated with the value are stored.

10. The server apparatus of claim 9, wherein the memcached logic to further receive a request to Update a second key in the memcached system with a second corresponding value,

determine whether the second key results in a hit, and

in response to the second key results in a hit, update a second status flag associated with a first item that yields the hit for the second key with a stale status, add the first item to a collect list, create a second item for the second key, and assign the second item with the second value.

11. The server of apparatus of claim 9, wherein the memcached logic to further periodically perform a maintenance operation to recover memory allocated to items having associated status flags with a stale status.

12. The server of apparatus of claim 9, wherein the memcached logic to further map physical addresses of the memory to an address space of the memcached system.

13. At least one computer-readable medium having instructions, to cause a computing device of a memcached system, in response to execution of the instructions by the computing device, to:

receive a request to Get, Set, or Update a value corresponding to a key in the memcached system;

determine, in response to the receipt of the request, whether the key results in a hit in a local cache maintained in memory of the client device; and

service, the Get, Set, or Update request based at least in part on whether a result of the determine indicates the key results in a hit in the local cache.

14. The computer-readable medium of claim 13, wherein the computing device is further caused to allocate storage spaces within the memory for the local cache.

15. The computer-readable medium of claim 13, wherein for a Get request with the key results in a hit, service comprises

- access the local cache to retrieve one or more addresses of one or more locations of the memcached system where the value and an associated status flag of the value are stored;

- perform a remote direct memory access operation on a server of the memcached system to retrieve the value and the associated status flag using the one or more addresses retrieved from the local cache; and

- determine whether the value retrieved through the remote direct memory access operation is stale, based at least in part on the status flag retrieved through the remote direct memory access operation.

16. The computer-readable medium of claim 15, wherein for a Get request with the key results in a hit, service further comprises return the retrieved value, in response to a result of the stale determination indicating the value is not stale.

17. The computer-readable medium of claim 15, wherein for a Get request with the key results in a hit, service further comprises:

- in response to a result of the stale determination indicating the value is stale,

- transmit the Get request to a server of the memcached system to request the value;

- receive second one or more addresses of second one or more locations of the memcached system where the value and an associated status flag of the value are stored;

- perform a remote direct memory access operation on a server of the memcached system to retrieve the value and the associated status flag using the second one or more addresses received;

- update the local cache to store second or more addresses for the key; and

- return the value received.

18. The computer-readable medium of claim 13, wherein for a Get request with the key not results in a hit, service comprises:

transmit the Get request to a server of the memcached system to request the value;
receive one or more addresses of one or more locations of the memcached system where the value and an associated status flag of the value is stored;

perform a remote direct memory access operation on a server of the memcached system to retrieve the value and the associated status flag using the one or more addresses received; and

return the value received.

19. The computer-readable medium of claim 13, wherein for a Set request, service comprises:

in response to the key results in a hit, invalidate a cache entry in the local cache that yielded the hit; and

send the Set request to a server of the memcached system.

20. The computer-readable medium of claim 15, wherein perform a remote direct memory access operation comprises first determine that a cache entry that yields the hit has not expired.

21. A method for memory caching data, comprising:

receiving, by a server of a memcached system, a request to Get an value corresponding to a key,

determining, by the server, whether the key results in a hit, and

in response to the key results in a hit, returning, by the server, one or more addresses of one or more locations of the memcached system where the value and a status flag associated with the value are stored.

22. The method of claim 21, further comprising:

receiving, by the server, a request to Update a second key in the memcached system with a second corresponding value;

determine, by the server, whether the second key results in a hit; and

in response to the second key results in a hit, the server, updating a second status flag associated with a first item that yields the hit for the second key with a stale status, adding the first item to a collect list, creating a second item for the second key, and assigning the second item with the second value.

23. The method of claim 21, further comprising periodically performing, by the server, a maintenance operation to recover memory allocated to items having associated status flags with a stale status.

24. The method of claim 21, further comprising mapping, by the server, physical addresses of memory of the server to an address space of the memcached system.