

61

of the particular type of instructions set, storage media, processor or processing strategy and may be performed by software, hardware, integrated circuits, firmware, micro code, or any type of other processor, operating alone or in combination. Likewise, processing strategies may include multiprocessing, multitasking, parallel processing and/or any other processing strategy known now or later discovered. In one embodiment, the instructions are stored on a removable media device for reading by local or remote systems. In other embodiments, the logic or instructions are stored in a remote location for transfer through a computer network or over telephone lines. In yet other embodiments, the logic or instructions are stored within a given computer, CPU, GPU, or system.

A second action may be said to be “in response to” a first action independent of whether the second action results directly or indirectly from the first action. The second action may occur at a substantially later time than the first action and still be in response to the first action. Similarly, the second action may be said to be in response to the first action even if intervening actions take place between the first action and the second action, and even if one or more of the intervening actions directly cause the second action to be performed. For example, a second action may be in response to a first action if the first action sets a flag and a third action later initiates the second action whenever the flag is set.

To clarify the use of and to hereby provide notice to the public, the phrases “at least one of <A>, , . . . and <N>” or “at least one of <A>, , . . . <N>, or combinations thereof” or “<A>, , . . . and/or <N>” are defined by the Applicant in the broadest sense, superseding any other implied definitions hereinbefore or hereinafter unless expressly asserted by the Applicant to the contrary, to mean one or more elements selected from the group comprising A, B, . . . and N. In other words, the phrases mean any combination of one or more of the elements A, B, . . . or N including any one element alone or the one element in combination with one or more of the other elements which may also include, in combination, additional elements not listed.

While various embodiments of the innovation have been described, it will be apparent to those of ordinary skill in the art that many more embodiments and implementations are possible within the scope of the innovation. Accordingly, the innovation is not to be restricted except in light of the attached claims and their equivalents.

What is claimed is:

1. A method comprising:

receiving a memory allocation request at a client device for primary memory from a component of the client device;

selecting, at the client device and in response to the memory allocation request, a subset of a region of memory in a memory appliance to be a portion of memory allocated at the client device, wherein the client device and the memory appliance are in communication over a network, and wherein the region of memory of the memory appliance is memory allocated for the client device before the memory allocation request is received;

mapping, at the client device, at least the portion of memory to an address space; and

accessing, by a hardware-accessible interface of the client device, data in the portion of memory via a client-side memory access, wherein a communication interface of the memory appliance is configured to access the subset

62

of the region of memory of the memory appliance as part of the client-side memory access.

2. The method of claim 1, wherein the hardware-accessible interface includes: a processor, a GPU, an MMU, an IO-MMU, a communication interface, an FPGA, an ASIC, a chipset, a hardware logic, and/or a memory access transaction translation logic.

3. The method of claim 1, wherein the hardware-accessible interface is accessed by a hardware application component.

4. The method of claim 1, wherein the hardware-accessible interface includes a memory and wherein the memory of the hardware-accessible interface is a cache for holding portions of the data.

5. The method of claim 1, wherein the hardware-accessible interface holds portions of the data in a portion of a memory of the client device.

6. The method of claim 1, wherein the hardware-accessible interface responds to and/or translates an attempt to access an address indicative of a location of the data.

7. The method of claim 6, wherein the hardware-accessible interface performs the client-side memory access in response to the attempt.

8. A client device comprising:

a processor;

a hardware-accessible interface configured to provide access to data using physical addresses; and

a local primary memory comprising:

instructions executable by the processor to select, at the client device and in response to receipt of a memory allocation request for primary memory, a subset of a region of memory in a memory appliance to be a portion of memory allocated at the client device, wherein the memory appliance is accessible from the client device over a network, and wherein the region of memory of the memory appliance is memory allocated for the client device before the receipt of the memory allocation request; and

instructions executable by the processor to map, at the client device, at least the portion of memory to an address space,

wherein the hardware-accessible interface is configured to provide access to data in the portion of memory via a client-side memory access, wherein a communication interface of the memory appliance is configured to access the subset of the region of memory of the memory appliance as part of the client-side memory access.

9. The client device of claim 8, comprising a hardware component from which the hardware-accessible interface is accessible.

10. The client device of claim 9, wherein the hardware component includes the processor, a GPU, a communication interface, a direct memory access controller, an FPGA, an ASIC, and/or a chipset.

11. The client device of claim 9, wherein the hardware-accessible interface is configured to participate in a cache coherency protocol with the hardware component.

12. The client device of claim 9, wherein the hardware component is configured to cache portions of the data in the local primary memory of the client device.

13. The client device of claim 9, wherein the hardware-accessible interface is configured to handle cache fill requests by a read of data from a memory and/or a cache included in the hardware component.

14. The client device of claim 8, wherein the hardware-accessible interface is configured to handle a cache fill

63

request by performance of the client-side memory access to read the data from the subset of the region of memory of the memory appliance, and/or wherein the hardware-accessible interface is configured to handle a cache flush request by performance of the client-side memory access to write the data to the subset of the region of memory of the memory appliance.

15 **15.** The client device of claim 8, wherein the hardware-accessible interface is configured to handle cache invalidate requests by an update to a memory and/or a cache of the hardware-accessible interface to indicate non-presence of data indicated by the cache invalidate requests.

16. A system comprising:

a client device; and

a memory appliance;

the client device comprising:

a processor;

a first communication interface;

a hardware-accessible interface configured to provide access to data using physical addresses; and

a local memory comprising:

an application logic unit; and

a client logic unit;

the memory appliance comprising:

a second communication interface;

a memory; and

a region access unit;

wherein the client logic unit is configured to cause allocation of a region of the memory that is accessible on the memory appliance by the client device through the first communication interface,

wherein the region access unit is configured to allocate the region of the memory for use as external primary memory of the client device,

64

wherein the client logic unit is further configured to: receive a request from the application logic unit to allocate a portion of memory; select, at the client device and in response to the request to allocate the portion of memory, a subset of the region of the memory in the memory appliance; and map, at the client device, at least the portion of memory to an address space addressable by the processor, wherein the hardware-accessible interface is configured to provide access to data in the portion of memory via a client-side memory access in which the second communication interface of the memory appliance is configured to access the region of the memory in the memory appliance as part of the client-side memory access.

17. The system of claim 16, wherein the hardware-accessible interface interfaces with a CPU interconnect.

18. The system of claim 16, further comprising a network including multiple communication paths between the client device and the memory appliance, wherein first communication interface of the client device includes a first plurality of communication interfaces, wherein the second communication interface of the memory appliance includes a second plurality of multiple communication interfaces.

19. The system of claim 16, further comprising a network over which the client device and the memory appliance are configured to communicate, wherein the network is a circuit switched network or a packet switched network.

20. The system of claim 16, wherein the client-side memory access is a Remote Direct Memory Access operation.

* * * * *