Claims

1. A method for managing migration of a source virtual machine in a data center network, wherein the data center network comprises a hierarchy of nodes, ~~A method for managing a data center network comprising a hierarchy of nodes,~~ wherein the data center network is configured to provide at least one service, the at least one service being associated with one or more virtual machines in the data center network, wherein requests and responses for the at least one service are respectively relayed downstream to and upstream from the one or more virtual machines through the nodes, each node comprising information for facilitating routing of the requests and responses through the nodes; and

wherein each node of the data center network comprises a Forwarding Information Base (FIB) with each FIB entry having:

a service name identifying a particular service associated with a virtual machine in the data center network; and

at least one interface associated with the service name, the at least one interface indicating an interface through which the virtual machine providing the particular service can be reached; and

wherein the data center network is managed according to a method comprising~~es~~:

(1a) sending a control message from a virtual machine through the nodes of the data center network, the control message indicating events related to the service associated with the virtual machine; and

(1b) at each node through which the control message is sent, updating the information based on the control message

wherein updating the information in (1b) comprises updating the FIB at the node, and

wherein during the migration, communication through the data center network is performed according to a method comprising the following steps at a node of the data center network:

(1c) receiving a *Request* packet comprising a request service name from an upstream node;

(1d) locating matching FIB entries having the service name matching the request service name; and

(1e) communicating with a further node based on the locating in step (1d) for further forwarding of the *Request* packet through the nodes.

~~2. A method according to claim 1, wherein each node of the data center network comprises a Forwarding Information Base (FIB) with each FIB entry having:~~
~~a service name identifying a particular service associated with a virtual machine in the data center network; and~~
~~at least one interface associated with the service name, the at least one interface indicating an interface through which the virtual machine providing the particular service can be reached; and~~
~~wherein updating the information in 1(b) comprises updating the FIB at the node.~~

2~~3~~. A method according to claim 1~~2~~, wherein the FIB entry further comprises a capacity value associated with the at least one interface, the capacity value indicating a number of requests that can be served per time unit by one or more virtual machines in the data center network associated with the service and reachable through the at least one interface.

~~4. A method for communicating through a data center network managed according to the method of claim 2 or 3, wherein the method comprises the following steps at a node of the data center network:~~
~~(4a) receiving a *Request* packet comprising a request service name from an upstream node;~~
~~(4b) locating matching FIB entries having the service name matching the request service name; and~~
~~(4c) communicating with a further node based on the locating in step (4b) for further forwarding of the *Request* packet through the nodes.~~

3~~5~~. A method according to claim 1~~4~~, wherein step (1e)~~4(c)~~ comprises:

if matching FIB entries are located, obtaining interfaces of the matching FIB entries as candidate interfaces;

if no matching FIB entry is located, obtaining upward interfaces of the node as candidate interfaces;

selecting the further node based on the candidate interfaces obtained; and

communicating with the selected further node.

46. A method according to claim 35, wherein selecting the further node comprises:

filtering tried interfaces from the candidate interfaces to obtain untried interfaces, the tried interfaces comprising interfaces to which the *Request* packet has been sent and from which a negative acknowledgement packet has been received; and

selecting the further node based on whether there are untried interfaces.

57. A method according to claim 46, wherein selecting the further node based on whether there are untried interfaces comprises:

if there are untried interfaces, selecting a node connected to an interface from the untried interfaces as the further node; and

if there is no untried interface, selecting the upstream node from which the *Request* packet is received as the further node.

68. A method according to claim 57, wherein selecting an interface from the untried interfaces comprises selecting the interface based on capacity values associated with the untried interfaces in the corresponding FIB entries.

79. A method according to claim 68, wherein selecting the interface based on capacity values associated with the untried interfaces in the corresponding FIB entries comprises distributing a request load of the service to each untried interface such that maximum utilization on the untried interfaces is minimized, wherein the utilization on each interface is the load on the interface per unit capacity of the interface.

810. A method according to any one of claims 57 – 79, wherein communicating with the further node comprises:

if there are untried interfaces, forwarding the *Request* packet to the further node; and

if there is no untried interface, sending a negative acknowledgement packet to the further node, the negative acknowledgement packet comprising contents of the *Request* packet.

9~~11~~. A method according to claim 8~~10~~, wherein each node of the data center network comprises a Pending Request Table (PRT), wherein the *Request* packet further comprises an indicator of a destination of the *Request* packet and wherein if there are untried interfaces, the method further comprises creating a PRT entry at the further node, the PRT entry comprising:

the request service name of the *Request* packet;

the indicator of the destination of the *Request* packet;

an incoming interface indicating the interface from which the *Request* packet is received; and

an outgoing interface indicating an interface to which the *Request* packet will be forwarded.

10~~12~~. A method according to claim 8~~10~~, wherein if there is no untried interface, the method further comprises updating the FIB at the further node based on the negative acknowledgement packet.

11~~13~~. A method according to claim 10~~12~~, wherein updating the FIB at the further node further comprises updating the interfaces of the FIB entries such that the interfaces associated with the request service name exclude the interface from which the negative acknowledgement packet is received.

12~~14~~. A method according to claim 11~~13~~, wherein the negative acknowledgement packet comprises the request service name and wherein updating the interfaces of the FIB entries comprises:

locating a matching FIB entry having a service name matching the request service name; and

if a matching FIB entry is located, removing from the interfaces of the matching FIB entry, the interface from which the negative acknowledgement packet is received.

13~~15~~. A method according to claim 12~~14~~, wherein updating the interfaces of the FIB entries further comprises:

if a matching FIB entry is not located, creating a new FIB entry at the further node, the new FIB entry comprising:

the request service name; and

interfaces comprising all upward interfaces of the further node, except the interface from which the negative acknowledgement packet is received.

14~~16~~. A method according to claim 13~~15~~, wherein the capacity values associated with the interfaces of the new FIB entry are equal.

15~~17~~. A method according to claim 13~~15~~ or 14~~16~~, further comprising converting the negative acknowledgement packet to a *Request* packet and forwarding the *Request* packet to one of the remaining interfaces of the matching FIB entry or one of the interfaces of the new FIB entry.

16~~18~~. A method according to any one of claims 9~~11~~ – 15~~17~~, further comprising:

receiving a *Response* packet comprising a service name and an indicator from a downstream node of the data center network, the indicator indicating a destination of the *Response* packet;

locating a matching PRT entry having a service name and an indicator respectively matching the service name and indicator of the *Response* packet; and

forwarding the *Response* packet to the incoming interface associated with the matching indicator of the matching PRT entry.

17~~19~~. A method according to claim 16~~18~~, further comprising removing the matching PRT entry if the *Response* packet is the last *Response* packet to be sent.

~~20. A method for managing migration of a source virtual machine in a data center network, wherein during the migration, communication through the data center network is performed according to a method of any one of claims 4 – 19.~~

18~~21~~. A method ~~for managing migration of a source virtual machine in a data center network, wherein the data center network is managed according to the method of~~

~~any one of claims 1 – 3 and wherein the method comprises~~ according to claim 1 or 2, further comprising:

sending a first control message from the source virtual machine through the nodes, the first control message indicating that the source virtual machine will stop providing a service;

copying memory pages from the source virtual machine to a destination virtual machine;

sending a second control message from the destination virtual machine through the nodes, the second control message indicating that the destination virtual machine will start providing the service.

19~~22~~. A method according to claim 18~~21 when dependent on claim 2~~, wherein each control message comprises a service name and wherein sending the first or second control message through the nodes further comprises the following steps at each node through which the control message is sent:

locating a matching FIB entry having a service name and an interface respectively matching the service name of the control message and the interface from which the control message is received; and

if a matching FIB entry is found, updating or removing the matching FIB entry based on the control message.

20~~23~~. A method according to claim 19~~22~~, wherein the control message comprises a capacity value and the source virtual machine is one of multiple virtual machines providing the service in the data center network, and wherein updating or removing the matching FIB entry based on the control message further comprises:

replacing the capacity value associated with the matching interface of the matching FIB entry with the capacity value of the control message.

21~~24~~. A method according to claim 19~~22~~ or 20~~23~~, wherein the method further comprises the following step at the node:

if a matching FIB entry is not found, creating a new FIB entry for the service.

22~~25~~. An apparatus comprising a data center network having a hierarchy of nodes,

wherein the data center network is configured to provide at least one service, the at least one service being associated with one or more virtual machines in the data center network, wherein requests and responses for the at least one service are respectively relayed downstream to and upstream from the one or more virtual machines through the nodes, each node comprising information for facilitating routing of the requests and responses through the nodes; and wherein ~~the data center network is managed according to a method of any one of claims 1 – 3~~ migration of a source virtual machine in the data center network is managed according to a method of claim 1 or 2.