Internal distribution code:

(A) [-] Publication in OJ
(B) [-] To Chairmen and Members
(C) [-] To Chairmen
(D) [X] No distribution

Datasheet for the decision of 21 November 2025

Case Number: T 1709/22 - 3.5.06

Application Number: 17758745.8

Publication Number: 3452909

IPC: G06F9/50

Language of the proceedings: EN

Title of invention:
SYSTEM AND METHOD FOR SUPPORTING FAST HYBRID RECONFIGURATION IN A HIGH PERFORMANCE COMPUTING ENVIRONMENT

Applicant:
Oracle International Corporation

Headword:
Sub-subnet/ORACLE

Relevant legal provisions:
EPC Art. 84, 56

Keyword:
Claims - clarity (yes)
Inventive step - (yes)

Decisions cited:

Catchword:

Europäisches Patentamt
European Patent Office
Office européen des brevets

Beschwerdekammern

Boards of Appeal

Chambres de recours

Boards of Appeal of the
European Patent Office
Richard-Reitzner-Allee 8
85540 Haar
GERMANY
Tel. +49 (0)89 2399-0

Case Number: T 1709/22 - 3.5.06

D E C I S I O N
of Technical Board of Appeal 3.5.06
of 21 November 2025

Appellant: (Applicant)	Oracle International Corporation 500 Oracle Parkway M/S 5OP7 Redwood Shores, California 94065 (US)
Representative:	D Young & Co LLP 3 Noble Street London EC2V 7BQ (GB)
Decision under appeal:	**Decision of the Examining Division of the European Patent Office posted on 14 February 2022 refusing European patent application No. 17758745.8 pursuant to Article 97(2) EPC.**

Composition of the Board:

Chairman	M. Müller
Members:	G. Zucka
	B. Müller

Summary of Facts and Submissions

I. The appeal is against the decision by the examining division, dispatched with reasons on 14 February 2022, to refuse European patent application 17758745.8, on the basis that neither the main nor the auxiliary request then on file satisfied the requirements of Article 56 EPC.

II. The following document is cited in the decision under appeal:

D2: US 2016/127236 A1.

III. A notice of appeal was received on 14 April 2022, the appeal fee being paid on the same day. A statement of grounds of appeal was received on 10 June 2022.

IV. The appellant requested that the decision under appeal be set aside and a patent granted on the basis of the claims of either request that was refused by the examining division.

V. The board issued a summons to oral proceedings. In the accompanying communication under Article 15(1) RPBA, the board set out its preliminary opinion, according to which the appealed decision should be upheld.

VI. In its response filed on 14 October 2025 with a letter dated 13 October 2025, the appellant requested that a patent be granted on the basis of the claims of the main request or one of the first to fourth auxiliary requests (the third and fourth auxiliary request

corresponding to the former main and auxiliary request, respectively), all filed with its response.

VII. During the oral proceedings before the board, the appellant filed a replacement auxiliary request 1 (hereinafter: auxiliary request 1), and withdrew the main request as well as the third and fourth auxiliary request.

VIII. The appellant thus requests that the decision under appeal be set aside and a patent be granted on the basis of the claims of the replacement auxiliary request 1 filed during the oral proceedings of 21 November 2025 or of the second auxiliary request filed with its letter dated 13 October 2025.

IX. The further text on file is:

description pages
1a and 2 to 27 received on 10 June 2022,
1 and 28 received on 20 November 2025;

drawing sheets
1 to 16 received on 10 June 2022.

X. Claim 1 of auxiliary request 1 reads as follows:

"A system for supporting fast hybrid reconfiguration in a high performance computing environment, comprising:
one or more microprocessors;
a first subnet (1300), the first subnet comprising a plurality of switches (501, 502, 503, 504), the plurality of switches comprising at least a leaf switch, wherein each of the plurality of switches comprising a plurality of switch ports and wherein each

of the plurality of switches is marked with a tuple that identifies its location,

a plurality of host channel adapters (510, 520, 530), each host channel adapter comprising at least one host channel adapter port,

a plurality of end nodes (1320), wherein each of the plurality of end nodes is connected to the first subnet via at least one host channel adapter of the plurality of host channel adapters;

wherein the plurality of switches of the first subnet are arranged in a network architecture having a plurality of levels including at least three levels, each of the plurality of levels comprising at least one switch of the plurality of switches;

wherein the system is configured to initially configure the plurality of switches according to a first configuration method for traffic across the whole of the first subnet, the first configuration method causing traffic to be routed in accordance with a first ordering of the plurality of end nodes;

wherein a subset of the plurality of end nodes are listed for reconfiguration;

wherein a subset of the plurality of switches is configured as a sub-subnet (1310) of the first subnet by comparing the tuples of leaf switches of the plurality of switches which are connected to the subset of the plurality of end nodes to select a common ancestor sub-tree comprising those leaf switches and their common ancestors, and marking all of the switches that belong to the common ancestor sub-tree for reconfiguration, the sub-subnet of the first subnet comprising a number of levels fewer than the plurality of levels of the first subnet;

wherein the system is further configured to reconfigure the sub-subnet of the first subnet according to a second configuration method that

reconfigures locally within the sub-subnet based on an internal traffic pattern of the sub-subnet only, wherein the leaf switches from among the marked switches are picked and the all of the end nodes connected to the picked leaf switches participate in the reconfiguration process, wherein the second configuration method comprises calculating a new set of routes only for the end nodes connected to the marked switches and distributing corresponding linear forwarding tables to the marked switches;

wherein the first subnet is a fat-tree topology subnet and wherein the sub-subnet of the first subnet is a sub-tree such that traffic within the sub-subnet of the first subnet is wholly contained such that traffic within the sub-subnet between the end nodes of the sub-subnet does not flow to or from the remainder of the subnet."

XI. Claims 7 and 12 of auxiliary request 1 relate to respectively a method and a computer readable storage medium having features corresponding to the apparatus features of claim 1 of that request.

XII. At the end of the oral proceedings, the chairman announced the board's decision.

Reasons for the Decision

1. *The application*

1.1 Although not claimed as such, the application is said to relate to the virtualisation of a high performance computer (HPC) system and live migration using an "SR-IOV vSwitch" architecture (description [0001]).

It is understood that "SR-IOV" stands for "Single Root I/O Virtualisation", i.e. a PCI Express (PCIe) specification that enables a physical device, e.g. a network interface card (NIC) to present itself as multiple virtual devices to a host system. The physical NIC (also called the Physical Function, PF) can expose multiple Virtual Functions (VFs), i.e. lightweight PCIe functions with dedicated resources (queues, registers, etc.) that can be assigned directly to individual virtual machines (VMs) or containers. Each VF provides near-native performance because traffic bypasses the hypervisor’s software data path and goes straight to the VM’s virtual interface.

A "vSwitch" ("virtual switch") is a software-based Ethernet switch running within the VM hypervisor. It directs traffic between VMs on the same host and between VMs and external networks. In a typical virtualised environment, all network I/O from VMs is processed by the vSwitch in software, then passed to the physical NIC. (This can cause extra overhead because every packet must traverse a software stack in the hypervisor.)

The reason for combining SR-IOV with a vSwitch is to achieve the performance gains of direct VF assignment (bypassing a software data path) while still retaining some of the benefits of a virtual switch (security, policy enforcement, QoS, etc.). An SR-IOV vSwitch architecture merges the speed advantages of SR-IOV with the centralised control and policy enforcement offered by virtual switches.

None of those features appear in the claims. They are merely mentioned to illustrate the context of the present application.

1.2 The aim of the application is to provide a fast reconfiguration of a High Performance Computing Environment. This is achieved by allowing for a partial reconfiguration of a subnet, thus avoiding a time-consuming full reconfiguration (description par. [0003]). The partial reconfiguration consists in limiting the reconfiguration to a subtree (sub-subnet) of the subnet where traffic between the end nodes is not flowing to of from the remainder of the subnet (see pars. [000107], [000112] and [000115] and figures 14 and 15).

2. *Auxiliary request 1 - clarity (Article 84 EPC) and interpretation of claims*

2.1 In claim 1 of auxiliary request 1, notwithstanding its name the highest-ranking request according to this decision, the board takes the view that the wording "system for supporting fast hybrid reconfiguration in a high performance computing environment" does not introduce - nor is intended to introduce - a technical

limitation but merely indicates the intended purpose of the claimed subject-matter.

2.2 In its summons (point 4.2), the board had observed that it was not clear in which sense an end node is "associated" with a host channel adapter.

This has now been clarified by specifying that the end nodes are connected to the first subnet via at least one host channel adapter.

2.3 The board had also observed (*ibid.*, point 4.3) that it was not clear what the "configuration" of the switches entails, what is the nature of the "first configuration method", what is the nature and the technical significance of the ordering of the end nodes, and in which sense the "first configuration method" is "associated" with this ordering.

This has now been clarified by specifying that "the system is configured to initially configure the plurality of switches according to a first configuration method for traffic across the whole of the first subnet, the first configuration method causing traffic to be routed in accordance with a first ordering of the plurality of end nodes".

The board accepts that the routing according to "a first ordering" may be considered a very broad but nonetheless clear feature. More specifically, the board accepts the appellant's argument made during the oral proceedings, that "the first ordering" is a parameter of the configuration algorithm which is irrelevant in the context of the claim (in particular for the technical effect on which the inventive step is argued).

The board also accepts the appellant's argument that the precise nature of the configuration and the reconfiguration methods is irrelevant, or whether the two methods are, in fact, different. What is relevant for the purpose of the claim is the scope of the configuration and reconfiguration methods, the configuration method operating in view of "traffic across the whole of the first subnet" and the reconfiguration method only operating "locally" in view of "an internal traffic pattern of the sub-subnet only", and the way the sub-subnet is defined on the basis of an initial selection of the end nodes which must participate in the reconfiguration process.

2.4 The board had further observed in its summons (*ibid.*, points 4.4 and 4.8) that, although claim 1 relates to a system, not a method, it was partially defined by method steps.

This has been clarified by specifying that it is the system which is configured to carry out such method steps. The board further holds that where this is not immediately apparent, viz. for the "marking" and the "picking" steps, the skilled person will also understand that the actor may be any suitable component within the claimed system.

2.5 The board had further observed (*ibid.*, point 4.4) that it was not clear what the "reconfiguration" of the sub-subnet entailed, and what was the nature of the "reconfiguration method".

This has been clarified by specifying that "the system is further configured to reconfigure the sub-subnet of the first subnet according to a second configuration method that reconfigures locally within the sub-subnet

based on an internal traffic pattern of the sub-subnet only".

2.6 The board had further observed (*ibid.*, point 4.6) that it was not clear in which sense traffic within the sub-subnet is "wholly contained".

This has been clarified by specifying that traffic within the sub-subnet between the end nodes of the sub-subnet does not flow to or from the remainder of the subnet.

2.7 The board had further observed (*ibid.*, point 4.9) that the selection of the "common ancestor sub-tree" was not clear.

This has been clarified by specifying that the common ancestor sub-tree comprises the leaf switches and their common ancestors.

2.8 As far as the nature of the "routes" is concerned, the board understands these to be the routes for the traffic mentioned earlier in the claim.

2.9 The "linear forwarding tables" are also understood as defining routes for said traffic.

2.10 The board, therefore, accepts that the appellant's amendments have overcome the board's objections regarding the clarity, Article 84 EPC, of claims 1, 7 and 12, of auxiliary request 1.

3. *Auxiliary request 1 - inventive step; Article 56 EPC*

3.1 The board is of the opinion that D2 is a suitable starting point for an inventive step analysis.

3.2 The system of D2 comprises one or more microprocessors (see claim 8).

It comprises a subnet (see par. [0020]).

The subnet comprises a plurality of switches comprising leaf switches (par. [0040]), each of the plurality of switches comprising a plurality of switch ports (par. [0042]).

The system comprises a plurality of host channel adapters (HCAs), each comprising at least one HCA port (figure 1; par. [0034]).

The system comprises a plurality of end nodes, each connected to the first subnet via at least one HCA (see figure 1).

The plurality of switches of the subnet are arranged in a network architecture having a plurality of levels, each level comprising at least one of the plurality of switches (see figures 2 to 8).

The plurality of switches is initially configured according to a first configuration method for traffic across the whole of the first subnet ([0021): "The subnet manager is responsible for configuring ... "; [0025): "The SM can calculate routing tables (i.e. the connections/routes between each pair of nodes within the tree) at network initialization time.")

The subnet is a "fat-tree topology subnet" (see par. [0018]).

3.3 According to the appellant,the following features of claim 1 are not disclosed in D2:

(a) The plurality of switches of the first subnet are arranged in a network architecture having a plurality of levels including at least three levels".

(b) A subset of the plurality of switches is configured as a sub-subnet of the subnet which comprises a number of levels fewer than that of the subnet, the configuration of the subset of switches as a sub-subnet taking place by comparing the tuples of leaf switches of the plurality of switches which are connected to the subset of the plurality of end nodes to select a common ancestor sub-tree comprising those leaf switches and their common ancestors, and marking all of the switches that belong to the common ancestor sub-tree for reconfiguration,

the system being further configured to reconfigure the sub-subnet according to a second configuration method that reconfigures locally within the sub-subnet based on an internal traffic pattern of the sub-subnet only, wherein the leaf switches from among the marked switches are picked and all of the end nodes connected to the picked leaf switches participate in the reconfiguration process, wherein the second configuration method comprises calculating a new set of routes only for the end nodes connected to the marked switches and distributing corresponding linear forwarding tables to the marked switches,

the sub-subnet being a sub-tree such that traffic within the sub-subnet is wholly contained such that traffic within the sub-subnet between the end nodes

of the sub-subnet does not flow to or from the remainder of the subnet.

3.4 The appellant agreed during the oral proceedings that feature (a) is at least implied as a possibility in D2 and would constitute an obvious choice for the skilled person. They further stated that the existence of at least three levels for the subnet in claim 1 merely intends to guarantee that the subnet is big enough to include a sub-subnet having enough levels (i.e. at least two) to constitute a sub-tree as defined in the claim: If the subnet only had two levels, it would have a single-leaf switch, which would render the claimed arrangement meaningless.

3.5 Feature (b) reduces the number of switches that need to be reconfigured, hence resulting in a reconfiguration that is substantially faster than the initial full configuration of the subnet.

3.6 The board holds that feature (b) is not disclosed in D2.

In particular, it holds that the "partition" in D2 cannot be mapped to the "sub-subnet" of claim 1, contrary to the reasoning given on page 4 of the appealed decision. In this respect, it accepts the appellant's point (appellant's response received on 14 October 2025, paragraphs spanning pages 5 and 6) that a "partition" in D2 is defined as a logical group of ports such that other members of the group can only communicate with other members of the same group (see D2, par. [0028]). Traffic between a subset of end nodes in D2 travels along paths defined at the port level or

using virtual lanes. D2 does not disclose the reservation of switches for exclusive use by a tenant.

As a consequence, D2 neither explicitly nor implicitly discloses a subset of switches that wholly contain traffic between a subset of end nodes, which would correspond to the "sub-subnet" of claim 1.

3.7 Furthermore, the skilled person would not have an incentive to introduce feature (b) in the system of D2. Instead, as set out in par. [0026] of D2, the skilled person would, to avoid reconfiguring the whole network, simply reconfigure those parts of the network where network changes occurred. This would, however, not be a sub-subnet, i.e. a sub-tree with traffic that is wholly contained as specified in claim 1. It would, instead, as set out in the appellant's response received on 14 October 2025 (paragraphs spanning pages 5 and 6) and as the appellant repeated during the oral proceedings, consist in a reconfiguration of a limited number of entire subnets, the reconfiguration effort further being reduced by segmenting a large fabric with routers.

In this respect, the board also accepts the point made by the appellant during the oral proceedings that in an Infiniband network, as in D2 (and implicitly in claim 1, since the presence of Host Channel Adapters implies an Infiniband network), a subnet is defined as being a set of hosts connected to a particular router which connects the subnet to other subnets (see D2, par. [0020]). To limit the scope of the reconfiguration, the skilled person would, therefore, attempt to optimise the arrangement of the routers.

3.8 According to the appealed decision (page 9, first paragraph), feature (b) merely expresses that all switches and end nodes affected by a desired change in sub-subnet configuration are identified and that the necessary changes are applied to their configuration information.

Such reasoning, however, firstly assumes that a "partition" in D2 can be mapped to a "sub-subnet" in claim 1, which is incorrect (see above).

Secondly, it fails to explain why the skilled person would come up with a "common ancestor", isolated sub-tree as in claim 1, with traffic between its end nodes that does not flow to or from the remainder of the subnet.

As it is, as explained by the appellant during the oral proceedings, the identification of such "common ancestor" sub-tree is a crucial aspect of the invention, which aims to reduce the reconfiguration effort after a hardware change in the network, e.g. after switching off and replacing an old switch by a new switch.

Switching off the old switch in a prior art system such as the one described in D2 will result in multiple link failures. No updated forwarding tables would at that moment exist for establishing new links after the replacement switch has been turned on. The impact of the change in the forwarding tables for each node cannot be known in advance. Therefore, to make certain that forwarding rules relating to the old switch, which may in principle be anywhere in the subnet, are corrected, it would be necessary to update all the

forwarding tables, i.e. to reconfigure the complete subnet.

To avoid such complete reconfiguration, the system of claim 1, in summary, identifies a sub-subnet which has its own routings, i.e. which has its own separate traffic between its end nodes, so that it is possible to change the forwarding rules purely within that sub-subnetwork, without affecting the rest of the subnet.

The board sees no reason why the skilled person, starting from the teaching of D2, would come up with the idea to limit the reconfiguration in this manner.

3.9 The board, therefore, arrives at the conclusion that the subject-matter of claim 1, and for analogous reasons claims 7 and 12, of auxiliary request 1 complies with the requirements of Article 56 EPC in view of document D2 alone.

4. Given that the examining division, in the decision under appeal, refused the application on this basis alone, without discussing further prior art, the Board is unable to remit the case with the order to grant a patent. It will rather be for the examining division to establish whether the requirements of Article 56 EPC are met in view of any other possibly relevant prior art.

Order

For these reasons it is decided that:

1. The decision under appeal is set aside.

2. The case is remitted to the examining division for further prosecution.

The Registrar: The Chairman:

L. Stridde M. Müller

Decision electronically authenticated