



(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
12.08.2015 Bulletin 2015/33

(51) Int Cl.:
H04L 12/803 ^(2013.01) **H04L 12/707** ^(2013.01)
H04L 12/721 ^(2013.01)

(21) Application number: **14162518.6**

(22) Date of filing: **29.03.2014**

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR
 Designated Extension States:
BA ME

(72) Inventors:
 • **Domzal, Jerzy**
30-059 Cracow (PL)
 • **Wójcik, Robert**
30-059 Kraków (PL)
 • **Dulinski, Zbigniew**
30-059 Kraków (PL)

(30) Priority: **05.08.2013 PL 40498613**

(74) Representative: **Fiolka, Janusz**
European Patent Attorney
Os. Dywizjonu 303, bl.21C/74
31873 Crakow (PL)

(71) Applicant: **Akademia Gorniczo-Hutnicza im. Stanisława Staszica w Krakowie**
30-059 Krakow (PL)

(54) **Device and Method for multiple path packet routing**

(57) The invention describes a device and a method for flow-aware multipath packet routing in computer networks. A router maintains a Flow Forwarding Table (FFT) containing identifiers of active flows and related router outgoing interface identifiers. Incoming packets which represent flows are analysed by the router. They are forwarded to the outgoing interface by an outgoing interface selector according to the Flow Forwarding Table (FFT), and the routing table is not consulted by the outgoing interface selector. When there is no entry corresponding to the flow represented by the incoming packet, the router adds the ID of the flow to the FFT, while the outgoing interface for the flow is taken from the current routing table. Accordingly, when the first packet representing a new flow arrives at the router, an outgoing interface identifier

is selected using existing methods utilising routing tables. Packets are then directed to an outgoing interface according to the FFT. When the cost of a link is increased by a routing protocol, for example due to congestion, the routing protocol propagates cost change information and recalculates routes using new costs. Accordingly, routing tables on routers may change. However, the FFTs remain unchanged. This means that only new flows with identifiers not present in the FFT during an update are affected by routing table changes. Active flows present on the FFT list during the update are unaffected by changes. In times of congestion new flows are directed along alternative paths, whereas all existing traffic travels via existing paths, without change.

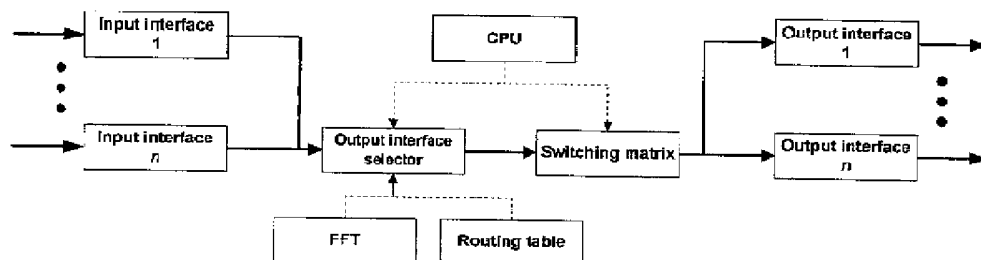


Fig. 1 Layout of a FAMTAR router

Description

[0001] The invention presents a device for multipath routing of packets in multipath IP (Internet Protocol) networks and the method for its use. The proposed solution allows for determination of transmission paths for traffic in an intelligent way, taking into consideration current link throughput levels. In this way, it is possible to allocate resources in a more effective way.

[0002] The main element of the device proposed by this invention is a router, which sets the paths for IP packets to destination nodes, and forwards packets to proper outgoing interfaces. In currently used routers, a routing table is used to determine the outgoing interface for the packet being served. The table contains information, stored in the router's memory, such as a list of nodes through which the destination networks are visible.

[0003] The goal of the present invention is to add a flow forwarding table (FFT) to the router, which will then handle and steer packets based on the FFT's content. The content of the FFT itself will be updated based on the routing table.

[0004] The routing table contains entries which allow for forwarding of the packets arriving at the router to each destination network or subnetwork. The routing table is set up and maintained in the physical memory of the router. A typical entry in the routing table contains the address of the destination subnetwork, a metric and an identifier or address for the outgoing interface, through which the subnetwork is available. In most cases, there is only one - the most favourable interface for each destination subnetwork in the routing table.

[0005] It is desirable that computer networks do not become congested, which is observed when more information than is possible to be sent needs to be processed. Load balancing is one method of eliminating congestion, the main goal of which is to send traffic to its destination node through several paths.

[0006] Load balancing allows for maintenance of several interfaces in the routing table with the same or different metrics, to the same destination subnetwork. Based on the current load, traffic is sent to each available outgoing interface according to assigned weights. In other words, the router sends packets to the destination network through different outgoing interfaces. Using load balancing, the packets of one flow may come to the destination node through different paths and in a different order. Moreover, this can cause the network to become inefficiently loaded. Some elements may become congested, and others not.

[0007] A group of packets may be considered to be 'a flow' when they are sent between the same nodes and have the same values in selected fields of their packet headers. These fields are usually the source and destination addresses, port numbers (source and destination) and the identifier for the transport layer protocol (in accordance with the Open System Interconnection - OSI/ISO model).

[0008] Currently, multipath transmission is possible under the MPLS (Multi-Protocol Label Switching) protocol. The MPLS standard is defined in RFC3031. Routers using the MPLS protocol send packets based on labels placed between the Layer 2 and Layer 3 headers. Altogether, routers using the MPLS protocol comprise the MPLS domain. A packet arriving at the border router of the MPLS domain, known as the Provider Edge router (PE router), is given the MPLS label and sent to the proper outgoing interface. Inside the MPLS domain, routers use only MPLS label for packet forwarding. Such routers are known as Provider routers (P routers). Each router on the packet's path has its own label switching table. Outgoing packets at the border of the MPLS domain have their MPLS label removed by the PE router, or one router before the PE router - outside the MPLS domain, the packet is served based on IP routing rules. Use of MPLS labels allows for the creation of un-equivalent paths and the subsequent sending of packets through different paths.

[0009] The label switching rules in routers must be set up before any packets are sent, as special protocols are used for label distribution in networks. The paths through which packets are sent also need to be set up beforehand. It is possible to distribute labels in such a way that the assigned paths are different from the paths estimated from the routing table, and this is an element of traffic engineering. Usually, paths which differ from those set up in the routing table have been set up statically by human operators, in most cases with Resource Reservation Protocol-Traffic Engineering (RSVP-TE), described in RFC3209 (updated and extended in RFC5151). MPLS does not allow for the creation and removal of paths dynamically, based on a network's current traffic load, but are set up permanently by the operator. MPLS also does not allow for alternative, optimal paths to be found in an operator's network. However, the router proposed in this present invention is able to find alternative, optimal paths.

[0010] The method by which a router selects the route to the destination node, wherein metrics like throughput, delay and jitter are considered, was presented in US Patent No. 61147, published in 2012. In it, several paths are established between two nodes, written in the routing table and periodically updated. As a result, the most effective path can be selected at any given moment.

[0011] An equivalent method for multipath packet routing in IP networks was presented in US Patent No. CN2011124418. Here, a central control system holding information on the entire network decides which path should be selected. In this way, the optimal path can be chosen based on the actual network conditions. The main advantages here are the minimisation of delays, increase of the amount of traffic transmitted through the network, and the limitation of jitter. The centralised control system, however, may result in lack of scalability.

[0012] The method for multipath transmission of packets aggregated into groups is described in US Patent No. WO2006EP65975. The aggregates are selected based

on a traffic matrix and are transmitted through paths selected by a central point in the network, with the goal of increasing the overall traffic transmitted. One of the drawbacks is, again, the central controller, which offers a focal point for failure.

[0013] The method for routing packets in mobile networks was proposed in US Patent No. 7242678, published in 2007. This patent, for something called "Edge Mobility Architecture (EMA)", demonstrated "Mobile Enhanced Routing (MER)", for packet forwarding in mobile networks. The signalling overhead generated as a result of a mobile device's IP address changes, is limited. This is made possible by generating unicast packages to update the mobile device's status between old- and new-access routers.

[0014] The method for packet routing in IP networks using multiple protocols in one domain was presented in US Patent No. 7177646, also published in 2007. This invention presents methods for using several types of routing updates in one domain.

[0015] The method for balanced routing of IP packets based on collection and analysis of traffic characteristics on selected transmission paths was presented in US Patent No. 7136357. The information collected is then distributed to other routers in the network, which then decide whether or not a path should be added to their routing tables. If there are several paths between the same nodes in the routing table, traffic is evenly distributed among them (each path has the same weight).

[0016] The method and system for routing of data streams in networks with multiple topologies were proposed in US Patent No. 8320277. In the proposed solution, resources are assigned proportionally according to weights connected with data streams. Moreover, the link metrics of particular topologies are also considered. For each topology, a separate routing table is maintained.

[0017] The method for intelligent routing of packets in Flow-Aware Networks was proposed in Polish Patent Application No. P.398761, which describes how packets are sent to selected outgoing interfaces based on the content of a protected flow list. The identifier of the outgoing interface is written to the protected flow list when the first packet of a flow arrives at the router. The outgoing interface is then selected based on the routing table

[0018] In the EIGRP protocol, described in detail in a Cisco document entitled "Enhanced interior Gateway Routing Protocol", ID=16406, several parameters are used to determine link costs. The most important are link bandwidth and delay. Current load in a link may also be observed. However, in such cases the costs of links can frequently change, and as a result the routing tables are updated more frequently. Allowing for creation of an increased number of loops in a network is also undesirable. However, the present patent is based on a mechanism which protects the network against loops.

[0019] The table of flow information, with instructions on how to treat them, was implemented in the OpenFlow switch (the 1.4 version of documentation for this switch

is available at www.opennetworking.org). The device presented in this document uses a similar table, albeit one that operates in a different way. In the OpenFlow switch, flow instructions are generated by the central controller, which is responsible for management of all devices in a network. The device presented herein works independently, creating its own new registrations in the flow table. In this way, the present invention operates in a similar way as the OpenFlow switch, however, the central controller is not needed. As a result, it allows for decentralised control, is cheaper, consumes less network resources, does not feature a single failure point, and is properly scalable.

[0020] The invention presented herein uses a new routing method designed for IP networks, FAMTAR (Flow-Aware Multi-Topology Adaptive Routing). Incoming packets which represent flows are analysed by the router. Next, they are forwarded to the outgoing interface according to the Flow Forwarding Table (FFT), which is a new element, unique to the submitted invention. When there is no entry corresponding to the flow represented by the incoming packet, the router adds the ID of the flow to the FFT, while the outgoing interface for the flow is taken from the current routing table.

[0021] FAMTAR uses the currently popular notion of identifying network traffic with the concept of flows. Although, the term 'flow' is already known from the literature, it is interpreted rather ambiguously. It is always, however, a stream of information belonging to one connection between two end users or applications. Flow ID can be established according to one of the established methods without impacting the present invention.

[0022] For example, the Flow-Aware Networking (FAN) architecture defines a 'flow' as a flight of packets, locatable in time and space and having the same unique identifier. The identifier is calculated as a hash of 5 header fields: IP addresses, transport layer port numbers and the ID of the transport layer protocol used (e.g. TCP or UDP).

[0023] The novelty in this invention is the method used to control flows in a network: for each packet, the outgoing interface is taken from the FFT (rather than from the routing table), with FFT entries added as soon as the packet of a new flow appears and which remain unchanged when the routing table changes.

[0024] A router is a network device which operates in the third layer of the OSI model and is used to connect computer networks - it plays the role of a switching node.

[0025] The submitted invention is illustrated in Fig. 1, which also presents an example usage.

[0026] Routers are composed of input interfaces, an output interface selector, a switching matrix, a routing table, a CPU and output interfaces. Each incoming packet is analysed. The output interface selector determines the proper outgoing interface based on the information from the routing table, and passes this information to the switching matrix. The switching matrix is responsible for physically forwarding the packet to the proper outgoing

interface. The operation of the outgoing port selector and the switching matrix is controlled by the CPU.

[0027] One element of every router is its routing table, which contains information on the interface that each packet should be forwarded to, based on the destination IP address found in the packet header. The routing table is therefore consulted for each incoming packet. When changes occur in the network (e.g. after failures, the addition of new networks, links, devices, routing policy changes, etc), the routing tables on all routers in the network are updated. Routing table updates impact all packets coming along after the update.

[0028] A new element in router architecture proposed in this invention and shown in Fig. 1 is the previously-mentioned Flow Forwarding Table (FFT). The FFT contains the outgoing interface identifiers on which packets belonging to given flows are forwarded, and is consulted by the outgoing port selector. Based on the flow identifier, the outgoing interface identifier is obtained. If the FFT contains information on a flow, the routing table is not consulted by the outgoing interface selector. This is a significant novelty in relation to state-of-art technique, because in current routers the routing tables are consulted for each and every packet. In the invention, if a certain flow is not present in the FFT, a new entry is created in the FFT, where the outgoing interface is taken from the current routing table.

[0029] Unlike the routing table, the FFT is static. Once created, the entries do not change, with one exception: the timestamp, which is the time at which the last packet of a flow appeared, is updated on each packet. Based on the timestamp, it is then possible to determine the elapsed time since the appearance of the last packet of each flow.

[0030] In the present invention, when congestion on a link occurs, the link cost is increased in the routing protocol, taking the maximum cost value or a special value indicating congestion. Consequently, applied routing protocol propagates cost change information and recalculates routes using new costs. Routing tables on routers may change, but in the invention, the FFTs remain unchanged. This means that only new flows with identifiers not present in the FFT during an update are affected by routing table changes. Active flows present on the FFT list during the update are unaffected by changes. Thanks to this behaviour, in times of congestion new flows are directed along alternative paths, whereas all existing traffic travels via existing paths, without change.

[0031] Figure 2 presents the structure of an FFT, which is stored in the physical memory of the router and contains at least the following fields:

- Flow identifier,
- Router outgoing interface identifier, through which packets belonging to respective flows are sent,
- Table entries enabling identification of precise time intervals between the arrival of the last served packet belonging to a flow and the arriving packet belonging

to this flow. If for a pre-defined period of time t a new packet belonging to the flow does not appear, its FFT entry is removed. If a packet belonging to the removed flow then appears after time t elapses, it is treated as a new flow.

The operating schema for an outgoing interface selector that a router uses is presented in Figure 3. A flow identifier is established based on the respective fields of IP packet headers. It is then determined whether or not this identifier is present in the FFT (**A**). If the flow identifier is present in the FFT, an outgoing interface number on a router is read (**B**) and the packet is directed to it. Next, the 'arrival time of the last packet' value is updated (**E**) (which can be represented by the current time), and the packet is sent to the interface indicated by the FFT using the switching matrix of the router (**F**). If the flow is not found in the FFT, then a respective outgoing interface is read from the router routing table (**C**). An identifier for this flow is added to the FFT (**D**) together with an outgoing interface number to which packets belonging to that flow should be directed. The rest of the procedure is similar to before, i.e. an 'arrival time of the last packet' value is updated (**E**) (current time may be registered) and the packet is sent to the interface indicated by the FFT using the switching matrix of router (**F**).

[0032] The essence of the proposed invention lies in the fact that for network operation and information transfer in a network it doesn't matter whether an entry in the FFT is placed first and a packet is then sent to an outgoing interface, or whether the packet is first sent to an outgoing interface and the next related entry is written to the FFT.

[0033] While, the order of information transfer and entry placement in the FFT is generally irrelevant, it is better to send a packet first and then add an entry to the FFT, and in this way minimise packet transfer delay.

[0034] With the IPv4 protocol, the Time to Live field (TTL) contains the figure for the maximum number of hops a packet can make on its path. Subsequent routers along the path decrease the value of the TTL field of each forwarded packet by 1. If a router gets a packet with a TTL equal to 0, the packet is dropped and removed from the network. This procedure helps avoid congestion when routing paths include misconfigured routers or when other failures occur. With IPv6, the field Hop Limit works identically.

[0035] When adding a new flow to the FFT, it is beneficial to store the TTL value from the IPv4 packet header or Hop Limit from the IPv6 packet header (depending on which protocol is used). Next, for each incoming packet, a router checks if the TTL (or Hop Limit) in the packet header matches the value stored in the FFT for the corresponding flow. If both values are the same, the packet is forwarded according to the procedure presented before. If the values do not match, the flow represented by the packet is erased from the FFT, and the packet processing procedure is restarted. In this way, after the TTL (or Hop Limit) mismatch is observed, a path for a

given flow may change, since the flow record was erased from the FFT (along with the recorded packet forwarding interface) and then recreated, but with an outgoing interface taken from the current routing table.

[0036] This represents another benefit of FAMTAR, as it eliminates the appearance of loops. Loops in FAMTAR may appear more often than in standard networks, because at the moment of the occurrence of a failure, or after any link cost change, routing protocols need some time to update the respective routing tables. Flows which appear in the network before changes happen are bound to proper interfaces, and without checking the TTL (or Hop Limit) field, those interfaces would never be changed. In this way, loops can occur which are not resolved by simple changes in the routing tables.

Claims

1. The patent requires a purpose-built router, in that the router is fitted with physical memory for storage of the FFT, which itself stores the flow identifiers and the router interfaces via which outbound traffic is sent.
2. Routers built according to Point 1 function in a specific way, in that the memory storing the FFT is accessed each time a packet enters the router, with the memory used to record the packet's outgoing interface.
3. The specificity of such routers also lies in the fact that if the FFT does not contain an entry for an analysed packet, then an entry for it is established from an existing routing table in the router.
4. Routers built according to Point 1 are also specific in that the memory storing the FFT contains at least two dedicated fields for the flow identifier and outbound interface identifier for packages belonging to this flow.
5. Routers built according to Point 4 are specific in that the FFT contains an additional field enabling time interval identification. This time interval expresses the time elapsed from the arrival of the last package in a given flow.
6. In the flow control method that is specific to this patent, router ports for a given outgoing packet are selected using the information stored in the FFT.
7. Routers built according to Point 1 are also specific in that the FFT includes an additional field for the value of the TTL field from the IPv4 protocol packet header, or the Hop Limit field from the IPv6 protocol packet header.
8. In the flow control method that is specific to this patent, if the TTL field value from the IPv4 protocol packet header (or the Hop Limit field value from the IPv6 protocol packet header) does not match the value stored in the FFT for the corresponding flow, that flow is removed from the FFT.
9. Routers built according to Point 1 are specific in that they analyse the traffic characteristics (for example, the load) of their outgoing links, and if these exceed a specified threshold, the router increases the cost in the routing protocol of the respective interfaces to a specified value, which then causes the distribution of the information to all other routers in the network and, consequently, prompts routing table changes in the subsequent routers.
10. Routers built according to Point 1 are specific in that routing table changes do not impact the paths of the ongoing flows.
11. Routers built according to Point 1 are specific in that routing table changes caused by traffic characteristic changes impacting link costs, do not change the paths of the ongoing flows.
12. The flow control method in a network according to Point 6 is specific in that routing table changes caused by traffic characteristic changes impacting link costs, do not change the paths of the ongoing flows.

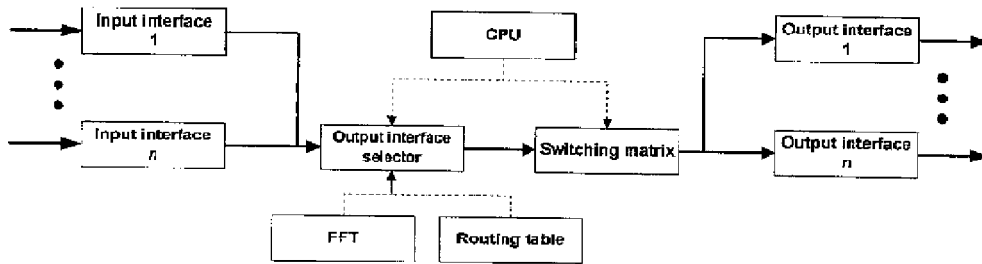


Fig. 1 Layout of a FAMTAR router

Flow ID	Outgoing interface ID	Timestamp (time of the last packet)

Fig. 2 Structure of the Flow Forwarding Table (FFT)

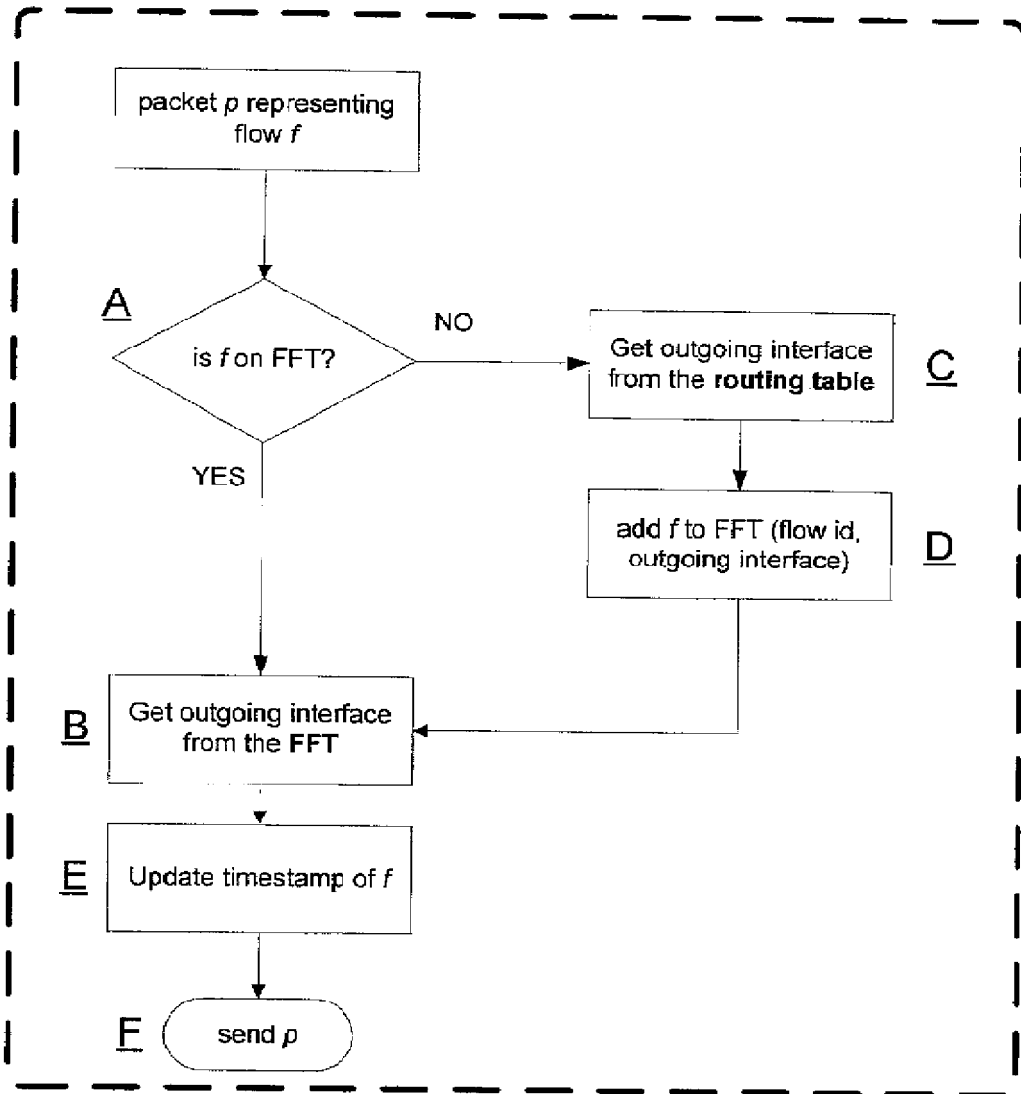


Fig. 3 Block diagram of the Output Interface Selector



EUROPEAN SEARCH REPORT

Application Number
EP 14 16 2518

5

10

15

20

25

30

35

40

45

50

55

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	US 2013/100957 A1 (SUZUKI HIDEAKI [JP] ET AL) 25 April 2013 (2013-04-25) * paragraphs [0074] - [0183]; figures 4-12 *	1-12	INV. H04L12/803 H04L12/707 H04L12/721
X	US 2004/243563 A1 (HEINER ANDREAS [FI] ET AL) 2 December 2004 (2004-12-02) * paragraphs [0004] - [0006] *	1-12	
X	US 2013/070762 A1 (ADAMS ROBERT EDWARD [US] ET AL) 21 March 2013 (2013-03-21) * paragraphs [0084] - [0091], [0096] - [0102]; figures 10D,12 *	1-12	
			TECHNICAL FIELDS SEARCHED (IPC)
			H04L
The present search report has been drawn up for all claims			
Place of search Munich		Date of completion of the search 1 July 2015	Examiner Köppl, Martin
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document	

1
EPO FORM 1503 03.02 (F04C01)

ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.

EP 14 16 2518

5

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

10

01-07-2015

15

20

25

30

35

40

45

50

55

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2013100957 A1	25-04-2013	CN 103081440 A	01-05-2013
		JP 2012009996 A	12-01-2012
		US 2013100957 A1	25-04-2013
		WO 2011162230 A1	29-12-2011

US 2004243563 A1	02-12-2004	EP 1423957 A1	02-06-2004
		US 2004243563 A1	02-12-2004
		WO 03021906 A1	13-03-2003

US 2013070762 A1	21-03-2013	AU 2012312587 A1	03-04-2014
		EP 2748974 A1	02-07-2014
		JP 2014531831 A	27-11-2014
		KR 20140060583 A	20-05-2014
		US 2013070762 A1	21-03-2013
		WO 2013043604 A1	28-03-2013

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- US 61147 A [0010]
- CN 2011124418 [0011]
- WO 2006 A [0012]
- EP 65975 A [0012]
- US 7242678 B [0013]
- US 7177646 B [0014]
- US 7136357 B [0015]
- US 8320277 B [0016]
- PL P398761 [0017]