Application of Rapid Spanning Tree Protocol for Automatic Hierarchical Address Assignment to Bridges

Dr. Arturo Azcorra, Universidad Carlos III de Madrid, Spain.
Eng. Guillermo Ibáñez, Universidad Carlos III de Madrid, Spain.

Abstract

This article describes a method to automatically assign hierarchical addresses in switched networks and proposes some applications ares to be further explored. Hierarchical addressing and routing are since long recognized as scalable and effective. However, its application in fixed and mobile ad hoc networks is restricted due to the overhead needed to group the nodes into clusters, maintain the hierarchy of nodes and the configuration complexity to assign and keep up to date the hierarchical addresses. The node differentiation due to the different roles at different hierarchy levels is also a disadvantage. The address assignment protocol uses the hierarchy set up by the Rapid Spanning Tree Protocol with the Designated Ports numbers of bridges as identifiers. A Bridge address results as the chain of designated port identities from the Root Bridge to that Bridge. These addresses express the node (bridge) position in the network. However, as IP and MAC addresses usage is well established in networks, the proposed hierarchical addresses shall neither tipically be used as layer two or layer three adresses, but just as routing aids. Other applications seem possible. Although this proposal is at a initial stage, application looks promising like enhancement of up/down routing protocol and Group leader election problems at network level. The protocol is simple and fully distributed among all spanning tree levels.

1 Introduction

Things are changing in Ethernet networks to cope with the challenges that campus and metropolitan area networks pose regarding configuration, bandwidth usage and network separation. Ethernet networks have evolved from the initial shared medium (the Ether) concept to dedicated point to point links due to performance and security reasons. To handle this growth in campus networks, complexity has been introduced in Bridges: Virtual LANs (VLANs), Multiple Spanning Tree (per VLAN spanning tree), link aggregation, IGMP snooping to handle multicast at layer two and other features. The configuration of some features may interfere with others, making the configuration of campus networks complex and critical.

There are proposals to reduce the complexity of Ethernet campus and metropolitan area networks. Our proposal is a generic hierarchical automatic addressing system for bridges (extensible to hosts at the cost of impact), that may be help in a set of different problems: Transparent Routing Bridges, Group Leader Election problems and other varied applications.

The assumption we make to use dedicated links in the connections between bridges is justified by the evolution from shared networks to switched networks, as it is reflected in the IEEE 802.1 standards. Use of dedicated links in campus networks is now recommended practice both by performance and security reasons. By performance to allow full duplex operation and by security to be able to implement access control to the network as specified by IEEE Standard 802.1X Port Based Access Control [2].This standard excludes shared ports as being non controllable.

This article is organized as follows : in section 2 we mention the related work in the areas related with our proposal: address assignment, transparent routing bridges, turn-prohibition protocols. In section 3 we describe the RSTAA protocol principle for fixed networks and in section 4 an overview of its potential applications .

2 Related Work

Spanning Tree Protocol use for labelling of nodes in sequential order has been proposed [18] for deadlock-free wormhole unicast routing. This simple ordered labelling of nodes (1,2,3,…) is used later for use by simplified protocols like up-down routing for deadlock prevention.

In [11], an idea for auto configuration of Source Routing Bridges (i.e. token-ring networks) uses the Root Bridge to assign number to bridges and LANs, a complex task for token ring networks. The Designated Bridge of each LAN asks for a number for that LAN. The Root Bridge must keep a table of the Designated bridges IDs, ports, LAN number. It is centralized on the root bridge and oriented to source routing bridges.

Regarding Automatic Address Assignment, a hierarchical IP address assignment based on Breadth First Search (BFS) tree construction is described in [14].

Group Leader election problems [17] at network level are also a related area where the author thinks that RSTAA might be applied, although it has not been yet analysed. Problem types in this area are: domain leader election in hierarchical networks, server assignment problem and multicast core assignment problem.

The subject of Bridges capable of routing is not new, it began with the source routing feature of token ring bridges, later evolved in Autonet [8] and later into Smartbridge [4]. The term Routing Bridge comes from the Source Routing Bridges used at Token Ring networks.

Rbridge is the term proposed by R. Perlman [6] for Routing Bridges. Spanning Tree Alternate Routing Protocol (STAR) [3] is another example of evolved Bridges that route between them. STAR bridges are compatible with 802.1D bridges. STAR proposed some concepts that Rbridges also employ [6] such as additional Layer 2 encapsulation and routing between enhanced bridges, but employs distance vector protocol instead of link state (IS-IS). STAR focuses on QoS looking for optimisation of network bandwidth allowing usage of direct links outside active topology between Star Bridges.

3 Rapid Spanning Tree Based Address Assignment (RSTAA) protocol

3.1 Rapid Spanning Tree Protocol

The extension of the Ethernet based switched domains based in campus networks (previously more segmented by routers or interconnected by ATM switches) makes the sometimes relegated Spanning Tree Protocol essential for network reliability.

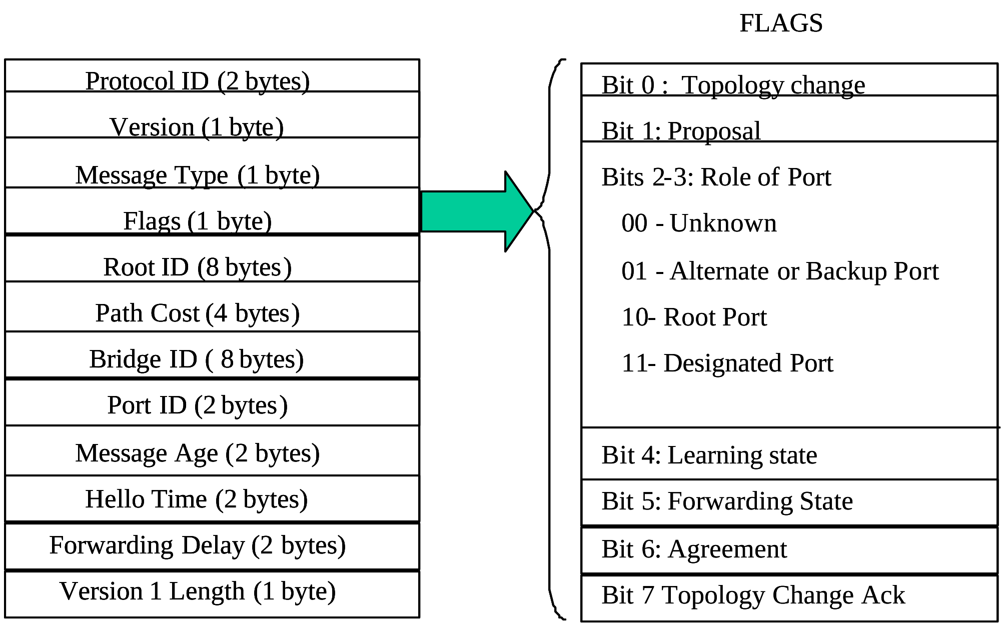
The relative long times for convergence of Spanning Tree Protocol, timer based protocol and the experience of several improvements (such as Port Fast Uplink Fast and Backbone Fast) have led to the optimised Rapid Spanning Tree protocol (IEEE 802.1w) that converges in times of 1-2 seconds instead of 30 seconds minimum for the Spanning Tree Protocol. Port states have been simplified (blocking, learning and forwarding) and port roles modified. From the timer based approach of STP it has been changed to the fast agreement principle between neighbour bridges. Special topology change notification BPDUs have been replaced by flags on the standard BPDU (figure 1) .

The propagation of topology changes is performed in one step all over the tree, instead the upward propagation to the root bridge first and then downward propagation of TCNs in STP. RSTP makes the addressing based on Spanning Tree very attractive due to its faster convergence and increased robustness. The intervals of unavailability due to rearranging of network in case of a topology change are greatly reduced .

3.2 RSTAA protocol

This address assignment system is based on the topology of the Spanning Tree. It is based on Rapid Spanning Tree Protocol and the relative connectivity information (BPDUs) interchanged between bridges using the Designated Port numbers downwards from Root Bridge as hierarchic coordinates. An example is shown in figure 2. The Root Bridge, the one with lowest Bridge ID (Bridge ID is an aggregation of the (configured) priority and bridge (MAC) identity), selected by running the new Rapid Spanning Tree Protocol (RSTP), is the origin of campus network link coordinates, the reference point for addresses, it has no coordinates. The assumption is that all

Fig. 1 Rapid Spanning Tree BPDU



links of campus network are point to point links with the exception of links to hosts (leaf nodes), which can be shared (unless the objective is to provide addresses to hosts instead to Bridges). In figure 2, Bridge 23 obtains this address because receives BPDUs from the Root Bridge via Designated Port of ID 23 of Root Bridge.

RSTAA BPDUs from neighbour RSTAA bridge containing its assigned RSTAA address.

Providing RSTAA addresses in a per host basis may be not practical in most cases due to the impact in host behavior needed to handle the RSTAA addressing and related routing protocol. The requirement for point to point links at all network levels including the hosts might also be too stringent for the installed based of hubs and shared links at host level.

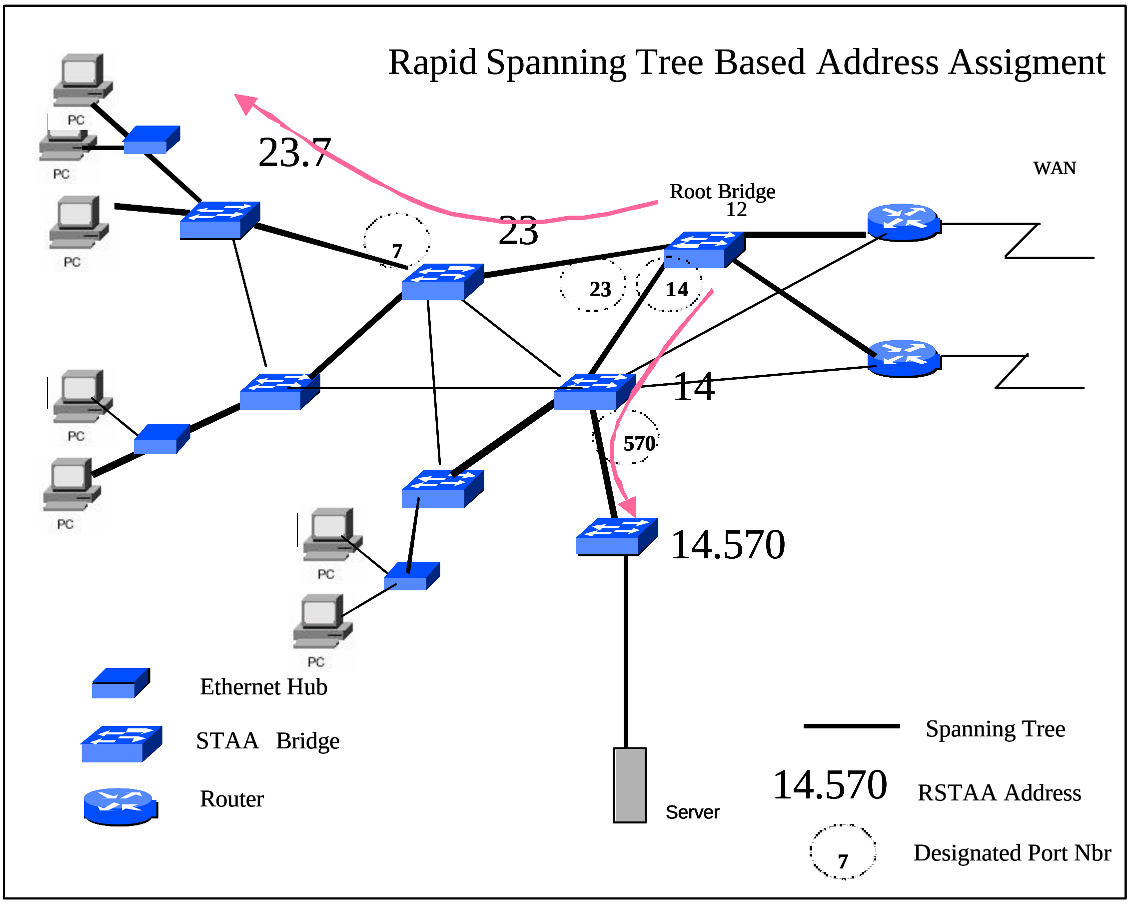


Fig.2 RSTAA addresses assignment principle

Bridge 23.7 obtains this address because it receives BPDUs from Root Bridge via the Designated Bridge 23 and its Designated Port 7.
RSTAA addresses identify topological position of a link. The address of the root port of each bridge is used as bridge topological “address” or coordinates .

3.2.1 Aplication of RSTAA to hosts

The example shown in figure 2 stops assigning addresses at bridge level. However, as long as point to point links are used, the addressing principle might also be applied at host level, where each host interface will get a unique RSTAA address. To do that, it is sufficient that hosts process

3.2.2 The RSTAA protocol BPDUs

The process of assigning RSTAA addresses to the Bridges takes place after Rapid Spanning Tree convergence. RSTP is reported to stabilize in 1-2 seconds. There are two implementation options for this Address Assignment protocol implementation: either extend RSTP with BPDUs for RSTAA addres assignment or implement as a separate protocol (RSTAA) running after RSTP converges and a Root Bridge is selected and the spanning tree is stable. For the rest of the section we will assume a separate protocol implementation, the RSTAA can access the information of RSTP status but is not integrated with it.

Making the RSTAA address assignment process a separate protocol that runs once RSTP is stabilized has some advantages regarding compatibility and independence from RSTP standard. In this case the process can operate as a waterfall. Routing Bridges, once RSTP stabilized, send new RSTAA BPDUs via their Designated Ports to their Designated Bridges. The BPDUs convey the full Bridge address (in a.b.c.d…. format) as shown in figure 3. The Root Bridge itself does not need strictly to send RSTAA Address Announcement BPDUs as there is no address to convey downwards the Spanning Tree. Anyway, Root Bridge will start the address assignment process by sending a BPDU with own address "0" to signal that he is the origin for RSTAA addresses, which is already known by receiving bridges as it is the Root Bridge. The Designated Bridges at first level below the Root Bridge send RSTAA BPDUs via their Designated Ports. The RSTAA BPDU can be like shown in figure 3. These BPDU contain their relative (RSTAA) address, which corresponds to the ID of Designated Port contained in RSTP BPDUs received from Root Bridge, which must be monitored to detect topology changes . In the example of figure 2 , Bridge labelled with 14 will send a “14” as its address downwards, besides its bridge ID. As the RSTP BPDUs carry the flag bits indicating it is a Designated Port, every Bridge can build its RSTAA Address from the RSTAA Address received at his root port from its designated bridge up in the Spanning Tree adding the Designated Port received in RSTP BPDU. Sending these BPDUs up via root port serves only as a confirmation of address assignment in upward direction. As the tree is stable during address assignment, the duration of address assignment can be very short.

Although it may be argued that there are enough addresses types with the existing IP and MAC, there are some clear advantages of RSTAA addresses: RSTAA addresses allow more compact routing tables than MAC addresses because can be consolidated, they contain topological information that can be used by hybrid or specific protocols. But due to the implantation of IP and MAC addresses can not be used in most cases as explicit addresses.

4 Potential Applications of RSTAA

RSTAA is an automatic address assignment protocol, its applications can be varied. The main advantage of RSTAA as an address assignment protocol is its automatism. The ways the principle can be applied are varied, for example if only the addressing features are needed, a hybrid RSTP+RSTAA protocol can be used only for address assignment, the forwarding of frames being completely handled by the routing protocols without broadcast via RSTP tree.

Protocol ID (2 bytes) (STAA)
Version (1 byte)
Message Type (1 byte) (Bridge Address)
Length (1 byte)
Root Bridge ID (8 bytes)
Designated Port ID 1st level (2 bytes)
Designated Port ID 2nd level (2 bytes)
Designated Port ID 3rd level (2 bytes)
....
Bridge ID

Fig.3 Example of RSTAA BPDU: RSTAA address announcement BPDU of a RSTAA Routing Bridge

4.1 Application to STAR (TRBs)

Using RSTAA topology information may help to implement hybrid protocols that use both the routes obtained from link state or distance vector interchanged with other nodes and the spanning tree paths.

Spanning Tree Alternate Routing protocol (STAR) [3] is a proposal for Transparent Routing Bridges that allows the use of alternative links besides those of the spanning tree. It is based on extension of STP to perform neighbour discovery of STAR bridges . STAR bridges can coexist with 802.1D bridges, improving the network bandwidth utilization, reducing average path length, thus contributing to QoS. STAR routing strategy is based on characterization of nodes according to its topological position whether a node is an ancestor, a descendent, or a node in a different branch of the tree. This information is easier to elaborate and more precise if the bridge transmits its RSTAA address. So it seems that an extension of STAR with RSTP and RSTAA is an interesting improvement.

4.2 RSTAA in Transparent Routing Bridges

As mentioned above, current campus networks have the following problems: large switching domains are prone to extensive frame storms, bandwidth is under utilized and costly infrastructure is inactivated by the STP. The administration of VLANs is complex and the associated Multiple Spanning Tree is also complex in big networks. Broadcasts shall also be reduced. The current objective for campus networks is to combine the plug an play features of transparent bridges with the network separation and

We can define Transparent Routing Bridges as routing devices that operate transparently to hosts, learning host location through received frames and routing them without the restrictions in topology and consequent bandwidth waste of the spanning tree algorithm. Each TRBs route transparently Layer 2 frames between them handling each one a list of parented hosts.

The use of RSTAA addresses in Adaptive Routing Bridges is introduced in [5]. However the approach that seems more compatible for operation with standard 802.1D bridges is to use these addresses *internally* as topological aids that are communicated by bridges via the routing protocols (for example adding them in Link State Advertisements) and *not as addresses* (i.e. explicitly included in the frame in address field). Transparent Routing Bridges use classic mechanisms as hello packets from routers with link state or distance vector advertisements that include the list of hosts parented by the announcing Bridge. A TRB becomes the Parent TRB for a number of hosts . The Parent Bridge should be the RSTAA TRB closest to the hosts (lower hierarchy RSTAA address). When one of these hosts starts communicating, the Parent Bridge will learn its MAC address and will include it its advertisements.

4.3 Application to location based protocols

RSTAA addresses are connection coordinates, then it seems that might be applied by location-based protocols, provided that a fixed relation of connectivity to location is maintained. This is normally a burden and does not provide any benefit, however in wireless networks such as 802.11 with infrastructure it seems that could be applied with advantage for routing. In the case of Ad Hoc networks [16], the RSTAA protocol needs a significant change to replace the role of the Designated Port in the address assigned with a substitute mechanism.

4.4 Application to wormhole routing

One of the most interesting and simple algorithms for deadlock prevention is the up/down routing [9][12][18]. Up/down routing has been applied to deadlock-free routing in irregular networks. It is simple but sub optimal as the shortest route may be prohibited by the up/down algorithm. Up/down routing is based in labelling all nodes according to a hierarchical tree and considers the links between the nodes as oriented (from the node with higher ID to the node with lower ID) It is the simplest of those based in prohibition of turns and is based in constructing an spanning tree only to label the nodes with numbers in sequence from the root bridge downwards. Only links at Spanning Tree are used in the local routing (simplest) form. RSTAA assigns instead true hierarchical addresses that express level in STP hierarchy. Instead of the basic comparison of node ID numbers to apply the TOP-DOWN restriction (turns (a, b, c) where b>c and b>a). Instead of this basic comparison, it seems possible to develop a more sophisticated up/down routing that might use transversal links outside the tree based on the comparison of nodes RSTAA addresses. For example, in up-down routing in local mode a legal path 9-3-4-6 if, using RSTAA the addresses of two of them not contiguous were 7.23 and 7.14, a "shortcut" link that does not use the spanning tree might be established maintaining the top-down restriction. Detailed analysis of this applicability is needed. Application of these coordinates to other simple routing algorithms for irregular networks seem also feasible.

3.5 Maintaining hierarchy and clusters

It is well-known that one of main disadvantage of hierarchical networks is the overhead to create and maintain the hierarchical structure. The other is the differentiation of nodes normally derived from the hierarchy.. Normally a clustering algorithm is needed to group the nodes into clusters and to elect cluster heads that disseminate the routes on behalf of cluster nodes to other clusters. Spanning Tree is a very efficient mechanism to elaborate a hierarchical tree with minimum communication cost in minimum time.

This applies not only to hierarchical networks but also to Group leader election problems, as described in [17]. Group Leader election problems [at network level are an area where RSTAA might be applied, although it has not been yet analysed. Problem types in this area are: domain leader election in hierarchical networks, server assignment problem and multicast core assignment problem

5 Conclusions

We have described a new protocol for bridged networks that uses the rapid spanning tree for assigning automatically unique and hierarchical addresses in a switched domain. These addresses are applicable to Transparent Routing Bridges, devices that intend to combine the plug and play characteristics of Transparent Bridges with the network separation provided by Routers. Many aspects have to be developed yet for a complete proposal.

Application to other network types and network problems such as wormhole routing, location based protocols, intra domain multicast and other group leader election problems in switched networks looks promising but needs further study as well.

6 Literature

[1] 802.1w - Rapid Reconfiguration of Spanning Tree. http://www.ieee802.org/1/pages/ 802.1w .html

[2] IEEE 802.1X Port based network access control. http://www.ieee802.org/1/pages/ 802.1x.html

[3] Lui K.; Lee W.C.; Nahrstedt, K.: STAR: a transparent spanning tree bridge protocol with alternate routing. ACM SIGCOMM Computer Communication Review, vol. 32, no. 3, pp.33-46, July 2002.

[4] Rodeheffer, T.L. ; Thekkath, C.A.; D.C. Anderson: SmartBridge: A Scalable Bridge Architecture. Proceedings of ACM SIGCOMM 2000, September 2000, Stockholm, Sweden.

[5] Azcorra, A.; Ibáñez, G.: Adaptive Routing Bridges: Nuevos enfoques para redes conmutadas. Conf. submitted to MUNDO INTERNET March 2004.

[6] Perlmann, R.; Williams, A: Design for a Routing Bridge. draft-perlman-zerouter-rbridge-00.txt . Internet draft. June 13, 2003.

[7] García, R. ; Duato, J.; Silla, F.: LSOM: A Link State Protocol over Mac Addresses for Metropolitan Backbones Using Optical Ethernet Switches. Proceedings of the second IEEE International Symposium on Network Computing and Applications (NCA'03)

[8] Shoreder, M. et al.: Autonet: A High-Speed, Self–Configuring Local Area Network Using Point–to–Point Links. IEEE Journal on Selected Areas in Communications, Vol. 9, No. 8, pp. 1318–1335, October 1991.

[9] Starobinski, D.; Karpovsky, G.; Zakrevsky, F.: Applications of Network Calculus to General Topologies, IEEE/ACM Transactions on Networking, June 2003, vol 11, No. 3, pp 411-422

[10] Perlman, R.: An Algorithm for Distributed Computation of a Spanning Tree in an Extended LAN, in the proceedings of Ninth ACM Data Communications Symposium, Vol. 20, No. 7, pp. 44–52, September 1985, New York, USA.

[11] Perlman, R.: Interconnections. Second Edition. Addison-Wesley, 2000.

[12] Zakrevski, L.; Jaiswal, Sharad; Karpovsky, M.: Unicast message routing in communication networks with Irregular topology. Proc. of CAD-99 1999.

[13] Kleinrock, L.; Kamoun, F. : Hierarchical routing for large networks, Computer Networks 1, 1977, pp. 155-174.

[14] Chamlee, M.A.; Zegura, E.W.; Mankin, A: Design and Evaluation of a Protocol for Automated Hierarchical Address Assignment, Proceedings Ninth International Conference on Computer Communications and Networks, pp. 328-333, 2000.

[15] Gigabit Campus networks Design. White Paper, Cisco Systems. 1999/2003. http://www.cisco.com.

[16]† Stojmenovic, I.: Position-Based Routing in Ad Hoc Networks. IEEE Communications Magazine, July 2002, pp.128-134

[17] Huang, Y.; McKinley, P.K.: Group Leader Election under Link-State Routing. IEEE O 1997 IEEE ISSN: 0-8186-8061-x/97

[18] Zakrevski, L., Karpovsky, M. "Fault-tolerant message routing in computer networks". Proc. Of Int. Conf. on PDP-A-99, 1999.