



Escola Politècnica Superior
de Castelldefels

UNIVERSITAT POLITÈCNICA DE CATALUNYA

PROJECTE DE FI DE CARRERA

TÍTOL: Millora de la gestió d'una xarxa WiFi: Guifi.net.

TITULACIÓ: Enginyeria de Telecomunicació (segon cicle)

AUTOR: Berenguer Vilajoliu Galceran

DIRECTOR: Roc Messeguer Pallarès

DATA: 23 de juny de 2006

Copyright (c) 2006 B. Vilajoliu.

Permission is granted to copy, distribute and/or modify this document under the terms of the GNU Free Documentation License, Version 1.2 or any later version published by the Free Software Foundation; with no Invariant Sections, no Front-Cover Texts, and no Back-Cover Texts. A copy of the license is included in the section entitled "GNU Free Documentation License".

Títol: Millora de la gestió d'una xarxa WiFi: Guifi.net.

Autor: Berenguer Vilajoliu Galceran

Director: Roc Messeguer Pallarès

Data: 23 de juny de 2006

Resum

Aquest estudi presenta la situació d'una associació d'usuaris anomenada Guifi.net. Els integrants d'aquesta comunitat han construït una xarxa WLAN oberta amb quasi set-cents punts de xarxa, una gestió dinàmica i una eina d'automatització per a nous nodes. Tanmateix, la confiança en les tecnologies basades en l'estàndard 802.11 ha implicat una expansió sorprenent el seu ús. Així doncs, el creixement de Guifi.net ha estat elevat, incidint en el rendiment d'alguns elements.

Com a conseqüència del creixement de la xarxa els usuaris estan comencen a experimentar una disminució del rendiment i disponibilitat. L'estudi que es presenta té com a principal objectiu solucionar aquesta problemàtica emergent a Guifi.net, a més de configurar una xarxa més robusta i escalable.

Durant la realització d'aquest projecte s'han testejat diversos protocols d'encaminament en escenaris concrets per veure el seu funcionament i extreure configuracions més eficients. L'entorn utilitzat en tot moment ha estat fidel a Guifi.net, des de la utilització del mateix hardware/software fins a les topologies utilitzades. Els coneixements adquirits s'han aplicat al cas concret de la xarxa oberta Guifi.net. Així doncs, es realitzen una sèrie de recomanacions per millorar la gestió de l'encaminament de la xarxa. Durant aquest estudi s'ha mostrat especial interès en el rendiment d'aquests protocols i tots els seus mecanismes d'escalabilitat, també aplicats al cas de Guifi.net.

Title: Management improvement in a WiFi network: Guifi.net.

Author: Berenguer Vilajoliu Galceran

Director: Roc Messeguer Pallarès

Date: June, 23th 2006

Overview

This thesis presents the actual situation of a community called Guifi.net. The members of this community have build an open WLAN network that includes almost seven hundred network points, with a dynamical management and a tool in which new nodes can be added automatically.

The growth of using technologies based on the standard 802.11 has implied a surprising expansion of it. In this way, the growth of Guifi.net has been high. This is the reason why the users are beginning to experience a diminution of the performance and availability.

This thesis is aimed at solving this emergent problematic in Guifi.net besides configuring a more robust and scalable network. During the accomplishment of this project diverse routing protocols have been tested in concrete scenarios to analyze their performance to extract more robust configurations. The environment used at all time has been true to Guifi.net, from the utilization of the same hardware and software besides the topologies used at the scenarios. The knowledge acquired during this thesis has been applied at the Guifi.net network. Therefore, some recommends are given in order to improve the management of network routing. During this study, the performance of routing protocols and their scalability mechanisms have been analyzed with special interest, always applied at Guifi.net.

ÍNDEX

INTRODUCCIÓ	1
CAPÍTOL 1. XARXA OBERTA: GUIFI.NET	3
1.1. Introducció	3
1.2. Serveis que ofereix Guifi.net	3
1.2.1. Serveis clàssics	3
1.2.2. Serveis multimèdia	4
1.2.3. Serveis de valor afegit i aplicatiu propi	4
1.3. La xarxa Guifi.net	5
1.4. Altres iniciatives a Catalunya	6
1.5. Altres iniciatives	6
CAPÍTOL 2. TECNOLOGIES UTILITZADES	7
2.1. Introducció	7
2.2. Hardware	7
2.2.1. Linksys WRT54GL	7
2.3. Software	8
2.3.1. OpenWRT	8
2.3.2. DD-WRT	8
2.4. Nivell d'enllaç	9
2.4.1. WLAN	9
2.5. Nivell de xarxa	9
2.5.1. OSPF	10
2.5.2. BGP	11
2.6. Propostes de Testbed	12
2.6.1. Encaminament	12
CAPÍTOL 3. PART PRIMERA: ENCAMINAMENT INTERIOR	13
3.1. Introducció	13
3.2. Primer escenari: OSPF i l'àrea troncal	14
3.3. Segon escenari: OSPF i particionat àrees	15
3.4. Tercer escenari: OSPF i sumarització de rutes	18
3.5. Quart escenari: OSPF i balanceig de càrrega	20
3.6. Cinquè escenari: OSPF extensiu	21
3.7. Consideracions de rendiment	23

3.8. Conclusions	25
CAPÍTOL 4. PART SEGONA: ENCAMINAMENT EXTERIOR.....	27
4.1. Introducció	27
4.2. Escenari 1: BGP com a protocol exterior (eBGP)	27
4.3. Escenari 2: BGP com a protocol interior (iBGP) i exterior.....	31
4.4. Escenari 3: BGP amb tres ASs.....	33
4.5. Escenari 4: extensibilitat de BGP	35
4.5.1. Reflectors de rutes	35
4.5.2. Confederacions	37
4.5.3. Servidors de rutes	38
4.6. Consideracions de rendiment.....	38
4.7. Conclusions	40
CAPÍTOL 5. APLICACIÓ A GUIFI.NET.....	41
5.1. Introducció	41
5.2. Disseny de la xarxa troncal	41
5.2.1. Disseny global de Guifi.net.....	42
5.2.2. Disseny de cada sistema autònom	44
5.3. Disseny de la xarxa de distribució	46
5.4. Implicacions d'implantació.....	48
CAPÍTOL 6. CONCLUSIONS FINALS	49
BIBLIOGRAFIA	51
ACRÒNIMS.....	53
ANNEX 1. PREPARACIÓ DELS WRT54GL	55
A1.1. Introducció	55
A1.1. Elecció de la distribució i versió.....	55
A1.1.1. OpenWRT.....	55
A1.1.2. DD-WRT	57
A1.1.3. Versions de Guifi.net	57
A1.2. Memòria persistent: NVRAM	57
A1.3. Failsafe mode.....	58
A1.4. El switch	58
A1.5. Metodologia de treball.....	60

A1.6. Bird	61
A1.7. Quagga	62
ANNEX 2. CONFIGURACIONS OSPF	63
A2.1. Introducció	63
A2.2. Escenari 2	63
A2.2.1. Router 4.....	64
A2.2.2. Router 2.....	67
A2.2.3. Router 3.....	69
A2.3. Escenari 3	73
A2.3.1. Router 4.....	73
A2.3.2. Router 3.....	75
A2.3.3. Router 2.....	77
A2.3.4. Test de tolerància a fallides.....	79
A2.4. Escenari 4	81
A2.4.1. Router 4.....	81
A2.4.2. Router 3.....	83
A2.4.3. Router 2.....	85
A2.4.4. Test de tolerància a fallides.....	86
A2.4.5. Test de balanceig de càrrega.....	87
A2.5. Escenari 5	88
A2.5.1. Router 4.....	88
A2.5.2. Router 3.....	90
A2.5.3. Router 2.....	92
ANNEX 3: CONFIGURACIONS BGP	95
A3.1. Introducció	95
A3.2. Escenari 1	95
A3.2.1. Router 4.....	96
A3.2.2. Router 3.....	98
A3.2.3. Router 2.....	101
A3.3. Escenari 2	103
A3.3.1. Router 4.....	103
A3.3.2. Router 3.....	106
A3.3.3. Router 2.....	108
A3.4. Escenari 3	110
A3.4.1. Router 4 (AS100).....	111
A3.4.2. Router 3 (AS300).....	113
A3.4.3. Router 2 (AS200).....	114
ANNEX 4: PLANIFICACIÓ I PRESSUPOST DEL PROJECTE	117
A4.1. Planificació	117
A4.2. Avaluació i pressupost	119
ANNEX 5: GNU FREE DOCUMENTATION LICENSE	121

INTRODUCCIÓ

Les possibilitats que ofereix la tecnologia de comunicacions de dades sense fils han facilitat que arreu s'hagin construït múltiples comunitats, independents als nous models de negoci sobre xarxes, i amb l'objectiu principal de connectar usuaris distants per així poder posar en comú recursos i serveis. Com que aquests no estan sota el control d'una empresa amb ànim de lucre, el seu llistat són el límit de les possibilitats tecnològiques, la imaginació i la capacitat de cada usuari.

Les comunitats Wi-Fi compleixen també una altra funció important. Aquesta és la de posar a l'abast de tothom les tecnologies tan aviat com algú estigui disposat a fer-ho i no a conveniència d'estratègies comercials de grans empreses, que no sempre actuen segons les condicions de mercat i possibilitats tecnològiques, ja sigui per posició dominant o en funció de la regulació existent.

Els beneficis que proporciona una xarxa amb aquest model són notables, ja que avui en dia la tecnologia crea excedents en recursos que no sempre podem aprofitar de forma particular, si es posen en comú no es renuncia a ells i a canvi s'obté accés als que els altres estan a la vegada disposats a compartir.

Guifi.net és una singular associació d'usuaris de xarxes sense fils a Catalunya. Aquesta és possiblement la que cobreix una extensió geogràfica major i possiblement la més activa, esdevenint tot un referent en el seu àmbit. Part dels usuaris mantenen una llista de correu on es discuteixen tots els temes de recerca i innovació que es duen a terme en aquesta associació. Tecnològicament, la xarxa de Guifi.net es podria definir com una maqueta de l'actual Internet, però amb menor quantitat de nodes i usuaris. Tanmateix, aquest no és un impediment per considerar Guifi.net com una xarxa IP en expansió, a la qual s'han d'aplicar els mateixos criteris de disseny i administració.

Aquest estudi és la materialització de l'inici de la col·laboració de Guifi.net amb el Departament d'Arquitectura de Computadors (DAC) de la Universitat Politècnica de Catalunya (UPC). D'entre els diversos fils de discussió a l'inici del present es considera molt important millorar la gestió de la seva xarxa IP. Donat que el fort creixement d'aquesta està implicant un empitjorament en la gestió dels nodes i adreces es creu interessant l'estudi i aplicació de diferents mecanismes que en puguin millorar-ne el rendiment. Posteriorment, i com a una de les tasques prioritàries d'aquest estudi és la redacció d'un seguit de propostes de millora per a Guifi.net.

El primer capítol d'aquest document realitza una introducció a l'entorn social on es desenvolupa aquest projecte. Així doncs, després de realitzar una presentació de Guifi.net es realitza una breu descripció d'aquesta comunitat, les seves aplicacions i característiques de la xarxa. Per acabar de plasmar un

estat de l'art de les xarxes obertes sense fils es fa una breu menció a altres iniciatives.

El segon capítol pretén realitzar un recull de les tecnologies a l'entorn de les quals es realitza aquest estudi, i que utilitza Guifi.net. Així doncs, es pot notar que aquest capítol ja suposa la primera part de desenvolupament d'aquest estudi, ja que també es cobreixen la descripció del maquinari disponible així com les diferents distribucions que s'hi poden instal·lar. Dins aquest capítol també s'aprofita anunciar les propostes que es desenvoluparan posteriorment i que resultaran en la resta de tasques d'aquest estudi.

El tercer capítol correspon a la segona part de desenvolupament d'aquest estudi. Dins aquest capítol es cobreixen tots els aspectes referents a la millora de l'encaminament interior del protocol IP. D'entre els diversos protocols existents se segueix amb l'estudi del protocol OSPF, ja que és el quin suporta actualment Guifi.net. Donada la disponibilitat d'equips hardware com els que s'utilitzen a Guifi.net es considera interessant el disseny, realització i optimització de diferents escenaris amb el protocol. Cadascun dels escenaris que es presenten dins d'aquest estudi cobreixen diferents problemàtiques o representen reproduccions de diferents zones de la xarxa de Guifi.net. Després de la realització d'aquest grup d'escenaris s'extreuen unes conclusions pròpies del protocol i que després s'aplicaran al cas concret de Guifi.net. Un altre punt molt important d'aquest estudi és l'avaluació del rendiment de cadascun dels protocols i per tant és en aquest punt on s'analitza el rendiment de l'OSPF.

El quart capítol implica la tercera part de desenvolupament. En aquesta secció es cobreixen totes les característiques de l'encaminament entre diferents dominis. Així doncs, s'utilitza el protocol BGP per a tal efecte. Anàlogament al punt anterior, després de decidir quines són les funcionalitats o problemàtiques concretes que es volen abordar, es realitzen una sèrie d'escenaris més o menys complexes per tal d'estudiar-les. El rendiment i l'extensibilitat del protocol són punts que s'analitzen amb detall. Finalment, i com en el capítol anterior, es realitzen una sèrie de conclusions del protocol, independents a l'entorn social dins del qual es realitza aquest estudi.

En el cinquè capítol representa una aplicació dels coneixements i conclusions extretes al cas particular de Guifi.net. Així doncs, es realitzen una sèrie de propostes d'implantació a Guifi.net. Donat que les tecnologies utilitzades s'estan executant actualment a Guifi.net aquesta proposta s'aproxima més a una proposta d'optimització, objectiu principal d'aquest estudi. Al final del capítol es realitza alguna consideració prèvia a la implantació de les recomanacions que es fan al llarg de tot el document i més especialment en aquest capítol.

Finalment, és sisè capítol recull les conclusions fruit de la realització de tot l'estudi i propostes d'optimització. Tanmateix, en aquest capítol també es cobreixen les línies futures i consideracions mediambientals del projecte.

CAPÍTOL 1. XARXA OBERTA: GUIFI.NET.

1.1. Introducció

L'objectiu principal de Guifi.net¹ és posar en comú la infraestructura dels usuaris i proporcionar els mecanismes d'organització per al seu funcionament i gestió. No és, en cap cas, una empresa ni privada ni governamental orientada a proporcionar un servei públic o de pagament, sinó simplement la coincidència en un interès comú del grup de persones que el forma. Guifi.net és, doncs, una comunitat amb una xarxa pròpia. Per formar-ne part no és necessari estar-hi connectat. La plataforma compta amb un portal que serveix d'espai de reunió, participació i divulgació.

Guifi.net també és un espai de recerca i desenvolupament per a l'adaptació de les noves tecnologies dins de les finalitats que li són pròpies en una xarxa oberta. Justament, fruit d'algunes col·laboracions i de la participació activa dels usuaris, la plataforma compta amb programari propi amb funcionalitats avançades i que a continuació és motiu d'estudi.

1.2. Serveis que ofereix Guifi.net

Guifi.net ofereix una sèrie de serveis als seus usuaris. A diferència d'un servei comercial aquests continguts estan administrats pels mateixos usuaris que els cedeixen i per tant, la seva disponibilitat o capacitat està limitada als recursos que aquest usuari pugui aportar. Tot i això, la cartera de serveis que aporta Guifi.net és més innovadora i avançada que qualsevol opció comercial per al gran públic.

1.2.1. Serveis clàssics

Tots els anomenats serveis clàssics de la pila TCP/IP estan implementats en diferents servidors, tots en l'entorn Linux. El parc de servidors distribuït aporta a la xarxa nom de domini registrat (guifi.net) i capacitat d'emmagatzematge de fitxers (FTP) per a tothom. A tots els usuaris que ho desitgin, Guifi.net configura una compta de correu electrònic del domini, accessible segons diferents mecanismes (POP/IMAP, SSL,...) i proporciona allotjament de pàgines d'usuari. En aquest punt, existeix una àmplia diversitat en quant a programari ja operatiu als diversos servidors, així com disponibilitat de pàgines dinàmiques i accés a bases de dades. També existeix un servidor de missatgeria en línia (IRC) que realitza funció de passarel·la amb la resta de canals existents a Internet.

¹ <http://www.guifi.net>

Per últim, alguns usuaris han habilitat portes a Internet i les cedeixen a qui les vulgui utilitzar. La llista està permanentment monitoritzada i publicada.

1.2.2. Serveis multimèdia

Al node bandoler.guifi.net, es troba en servei de veu sobre IP (VoIP). El node disposa d'un servidor Asterisk. El servei és accessible des de Guifi.net, i també via NAT i Internet. Compta amb bústies de veu, amb avís de missatges via SMS i correu electrònic. Tothom hi pot crear extensions. En l'actualitat existeixen serveis de pagament semblants, com és el cas de iPhone d'Al-Pi Telecomunicacions, però els grans operadors actualment no ofereixen cap tipus de suport ni servei semblant.

1.2.3. Serveis de valor afegit i aplicatiu propi

Des del portal públic, punt principal de trobada i difusió, hi ha accessible un ventall d'aplicacions per a tot grup en xarxa: blog, fòrum, etc. Des de la mateixa interfície es pot accedir a l'aplicatiu propi que es desenvolupa a la plataforma. D'entre aquests és interessant nomenar-ne dos, els quals són alguns dels punts forts i referents de Guifi.net:

- **MapServer:** Servidor de Mapes. Fruit del contacte amb l'Institut Català de Cartografia s'ha integrat l'aplicació d'aquesta entitat per tal de poder utilitzar els mapes que aquesta entitat cedeixen. A Guifi.net aquests mapes serveixen com capa de l'aplicació que mostra l'estat dels nodes i els enllaços de la xarxa. Gràcies a aquesta col·laboració, la informació dels nodes està georeferenciada. El fet de poder aprofitar l'eina Minnesota Map Server ha minimitzat els costos a l'hora d'implementar-la i implantar-la. L'aplicació MapServer esta disponible per a tots els usuaris que la vulguin utilitzar, a més que s'utilitza com a un dels mòduls de la següent eina que es presenta.
- **UnSolClick:** Eina d'autoconnexió a Guifi.net. Qualsevol usuari que vulgui pertànyer a la comunitat ha de registrar-se al portal col·laboratiu. Si a més, l'usuari vol connectar-se a la xarxa pot utilitzar l'eina de configuració i connexió automàtica. Un cop s'ha introduït les coordenades es localitza l'usuari al servidor de mapes i s'obté una llista de possibles nodes d'interconnexió. L'usuari rep un plànol amb la seva situació i la dels possibles nodes amb els que pot enllaçar. Per a cadascun dels nodes, es disposa de més informació (perfils de relleu, per exemple) que l'usuari també pot consultar per escollir de forma idònia amb quin node s'enllaçarà. Un cop seleccionat el node el servidor retorna la configuració de xarxa que ha d'introduir l'usuari al seu punt d'accés i adreçament per als seus dispositius. Cal dir que aquesta informació no es limita a adreça, màscara i porta d'enllaç IP, sinó que es tracta del fitxer de configuració per al sistema operatiu utilitzat, i amb tota la configuració avançada que permet una xarxa amb aquest model.

1.3. La xarxa Guifi.net

La xarxa s'estén a diverses comarques catalanes, generalment a Girona i Barcelona, encara que una de les seves característiques diferencials recau en la seva extensibilitat: la xarxa Guifi.net la construeixen els mateixos usuaris així que a priori, no hi ha limitació geogràfica. A l'hora de desplegar la xarxa Guifi.net es segueixen una sèrie de criteris comuns i es configuren els protocols OSPF i BGP. Aquest últim està encara en proves i no aplicat a tot l'equipament.

La topologia de xarxa de guifi.net és una gran malla (mesh) que alhora està formada per dos grans parts. La primera és part troncal que s'encarrega de intercomunicar les diferents zones, i els "punts guifi.net" o nodes, des d'on es proporciona cobertura als clients. La part troncal són enllaços típicament configurats com a enllaços entre dos punts i gestionats de forma comuna entre tots els usuaris.

A diferència d'una xarxa convencional en estrella, basada en troncsals en suport físic i d'alta capacitat, en les comunicacions sense fils la part troncal no disposa de més capacitat que la part d'accés o el mateix punt final (xarxa local de l'usuari). Per tant, és important dispersar el tràfic a través de diversos enllaços troncsals per augmentar la capacitat total de la xarxa.

Els "punts calents", o "punts guifi.net" ja són els nodes que cada usuari munta i estan gestionats lliurement per qui es vulgui fer responsable d'aquesta tasca. Un node es munta amb la tecnologia que cadascú tingui disponible al seu abast i vulgui amb la única restricció de no perjudicar el bon funcionament de la xarxa troncal. Accepta connexions lliurement de qualsevol client o no en funció dels criteris de qui els administren.

Les adreces s'assignen seguint la política de Freenetworks.org. Aquesta iniciativa pretén assignar adreçament privat d'una forma jeràrquica i topològicament vàlida a tots els clients que s'hi vulguin adherir. L'única condició és permetre el lliure trànsit, com si d'una rèplica de l'actual Internet es tractés. L'aplicació anirà assignant de forma automàtica les adreces als usuaris i enllaços d'acord amb els rangs que assignin els administradors de cada zona. A mode d'exemple, el criteri és el següent:

- Xarxes /27 (32 adreces, màscara 255.255.255.224) per als punts d'accés. Per exemple dins del rang 10.138.0.0/15
- Xarxes /30 (2 adreces, màscara 255.255.255.252) per a les connexions punt a punt dels troncsals (amb algunes excepcions). Per exemple dins del rang 172.25.0.0/16
- 192.168.0.0/16 queda lliure per muntar-se subxarxes particulars.

Els llistats de nodes i enllaços s'obtenen amb pàgines dinàmiques generades a partir de la informació que hi ha a la base de dades de l'aplicatiu de Guifi.net.

1.4. Altres iniciatives a Catalunya

La iniciativa Guifi.net no és ni única ni pionera a Catalunya, ja que actualment n'existeixen fins a un centenar, que cobreixen gran part del territori. Tots aquests grups de treball convergeixen sota uns objectius i metodologies semblants. Des de fa temps totes aquestes organitzacions s'agrupen per realitzar trobades i posar en comú algunes experiències i coneixements. El punt comú de totes aquestes iniciatives és que s'adhereixen a un text, anomenat "The Commons Wireless Manifesto", i que detalla el model a seguir per a que totes aquestes xarxes lliures puguin existir i créixer de forma sostenible.

1.5. Altres iniciatives

Per completar un estat de l'art de les xarxes lliures sense fils s'ha de comentar que aquest moviment no és únic ni està subjecte a classes o distincions arreu del planeta. Dins l'estat espanyol també existeixen moltes comunitats obertes. És considera important nomenar la comunitat que s'ha muntat a Guadalajara, sota el nom de Guadawireless². Aquesta iniciativa segueix el mateix model i eines tecnològiques que Guifi.net. La seva xarxa compta amb els mateixos serveis, als que cal afegir la missatgeria instantània i jocs en xarxa. Com a característica semblant, Guadawireless també compta amb un sistema per georeferenciar els nodes i mostrar-ho de forma dinàmica amb el portal propi.

Possiblement el moviment més semblant a Guifi.net, amb el mateix model característic però una extensió molt major és AthensWireless³. Aquesta iniciativa compta amb més de tres-mil nodes dins de la seva xarxa i els mateixos serveis. També són molt conegudes les comunitats de Seattle, New York, Paris i Berlín.

Paral·lelament a l'aparició de tant i tant variades comunitats sense fils han sorgit altres comunitats virtuals, també anomenades metacomunitats. Una de les funcions socials que absorbeixen són la comunicació i integració entre les diferents comunitats existents. Les més conegudes són Freenetworks, Freifunk i Redlibre, la primera ja comentada en aquest capítol ja que Guifi.net segueix el seu pla d'adreçament.

Totes les comunitats nombrades fins al moment tenen en comú que son lliures i impulsades pels usuaris. Tanmateix, també existeixen altres moviments promoguts amb ajuts públics o privats. Per exemple, ho és tot ajuntament que instal·la una xarxa sense fils al municipi i l'ofereix a baix cost als seus usuaris. La resta d'opcions per obtenir accés sense fins des de la llar ja es poden etiquetar de comercials i per tant, treball no pretén estudiar.

² <http://www.guadawireless.net>

³ <http://www.athenswireless.net>

CAPÍTOL 2. TECNOLOGIES UTILITZADES

2.1. Introducció

Un cop explicat l'entorn social en el que es realitza aquest estudi cal veure en quin entorn tecnològic es desenvolupa. Aquest punt és necessari a fi de conèixer de quines eines (hardware i software) es disposa per realitzar aquest estudi i exposar en quines tecnologies es vol centrar. Aquest capítol també pretén presentar aquestes tecnologies i nomenar els coneixements mínims necessaris per continuar la lectura del següent capítol. Per complementar la informació, el lector es pot adreçar a l'annex primer, titulat "Preparació dels WRT54GL".

Per últim, és considera important la redacció d'aquest punt a fi d'obtenir una llista de propostes realitzables en quant a l'entorn tecnològic i la seva disponibilitat.

2.2. Hardware

La xarxa Guifi.net es basa en un equipament comú per formar la part troncal i en l'equipament propi, per als punts finals. Però la topologia altament mallada comporta que gran part de la infraestructura siguin els mateixos equips dels usuaris. Així doncs, la xarxa està formada per una quantitat elevada de punts d'accés domèstics. Cal dir, que una part important d'aquests equips son del fabricant Linksys.

Per a la realització d'aquest estudi es disposa de tres punts d'accés inal·làmbrics del mateix fabricant, Linksys, filial corporativa de Cisco Systems.

2.2.1. Linksys WRT54GL

Després de l'evolució dels models WRT54GS i G el fabricant ha apostat per el nou model GL compatible amb altres distribucions basades en Linux, ja que les últimes versions (WRT54G v.5) dels primers ja no el suporten. En quant a diferències hardware, el nou model manté la mateixa quantitat de memòria RAM i NVRAM que els primers de la saga, mentre que els nous models GS i G l'han reduït.

Aquest equip és un encaminador amb tres interfícies de xarxa: la primera per enllaçar amb l'operador (WAN), la segona correspon als quatre ports de xarxa Fast Ethernet (LAN) i la tercera és la interfície ràdio (WLAN). La seva potència de transmissió és de fins a 18dBm. L'equip és compatible amb els estàndards 802.11, 802.11b, 802.11g i 802.11u.

Els models en concret del qual es disposa no compta amb ports RS-232, que faciliten la instal·lació de nou programari o firmware.

2.3. Software

En quant a software es considera el firmware original que incorpora l'equipament i algunes versions de totes les disponibles i compatibles amb el hardware disponible. És important escollir aquelles distribucions que ofereixen més llibertat d'operació i disponibilitat d'altres eines software funcionals, ja vinguin de sèrie o no. Per tant, i d'entre totes les distribucions possibles se'n consideren dues basades en Linux i amb la màxima disponibilitat d'aplicacions i llibreries software: OpenWRT i DD-WRT.

2.3.1. OpenWRT

La distribució OpenWRT⁴ està especialment pensada per a encaminadors i distribuïda sota llicència GPL. En comptes d'introduir el màxim de funcionalitats possibles en la distribució només es proporciona un firmware amb el mínim de capacitats. Per als usuaris amb coneixements avançats de Linux hi ha suport per afegir una varietat molt ampla de paquets i programari, afegint les funcionalitats que es desitgen de forma modular.

Com que OpenWRT es tracta d'una distribució oberta, els usuaris desenvolupen o adapten altres programaris existents. Així doncs, OpenWRT és la distribució amb el màxim de funcionalitats disponibles: dimonis d'encaminament, VLANs, VoIP, encriptació, seguretat i una extensa llista.

Per últim, comentar que aquesta distribució també és compatible amb equipament de D-Link, Asus, Dell, Buffalo, Motorola, Toshiba, US-Robotics i alguns més.

2.3.2. DD-WRT

DD-WRT⁵ és la segona distribució més versàtil de les compatibles amb el Linksys WRT54GL. Es tracta d'un programari totalment lliure i distribuït sota llicència GPL.

L'última versió de DD-WRT, anomenada v23, és un projecte totalment nou respecte la primera iniciativa. Les funcionalitats són tant o més amples que en el cas de l'OpenWRT. És d'especial interès nombrar-ne algunes de diferencials com el suport inicial per IPv6, QoS i una interfície Web amigable.

Per últim, i igual que en l'anterior distribució, comentar que DD-WRT també és compatible amb equipament d'Asus, Siemens, Motorola, Buffalo i Allnet.

⁴ <http://openwrt.org>

⁵ <http://www.dd-wrt.org>

2.4. Nivell d'enllaç

La capa d'enllaç de dades és la capa encarregada de transferir dades entre nodes de xarxa adjacents en una WAN o entre nodes del mateix segment de LAN. La capa d'enllaç de dades proporciona les funcions i els procediments per a la transmissió de dades entre entitats de xarxa i pot detectar i, fins i tot, corregir errors que poden ocórrer a la capa física. Exemples de protocols d'enllaç de dades són Ethernet per a xarxes d'àrea local i PPP o HDLC per a connexions punt a punt. L'enllaç de dades proporciona transferència de dades al llarg de l'enllaç físic. Aquesta transferència pot ser o no fiable; molts protocols d'enllaç de dades no disposen de reconeixement de recepció i acceptació de trames amb èxit i, fins i tot, alguns protocols ni tant sols disposen de checksum per tal de comprovar errors en la transmissió. En aquests casos, protocols de més alt nivell han d'encarregar-se de proporcionar control de flux, comprovació d'errors, reconeixement i retransmissió.

2.4.1. WLAN

La tecnologia WLAN és regida per l'estàndard 802.11, especificat per l'Institute of Electrical and Electronics Engineers (IEEE), el mateix organisme encarregat d'especificar la resta de tecnologies de xarxa (com l'ethernet). Mitjançant aquest procés, es garanteix inicialment la interoperabilitat entre diferents fabricants i un funcionament dins les capes del model de comunicacions igual que qualsevol altra tecnologia de LAN. Per tant, la seva interacció amb protocols de comunicació, com TCP/IP, és totalment transparent. Des del punt de vista del sistema operatiu, emprar un adaptador de WLAN és igual que fer-ne servir un de LAN tradicional.

Aquest projecte no estudia aspectes ràdio ni del canal, però si es considera important aquesta breu menció, ja que pel fet d'utilitzar una tecnologia sense fils i un medi hostil, els enllaços poden tenir una disponibilitat variable. Aquesta condició de contorn pot afectar a les capes superiors (nivell de xarxa) d'una forma diferent a per la qual han estat ideats. A més, es considera interessant aquesta menció, donat que les tecnologies basades en 802.11 utilitzen una banda lliure de l'espectre i per tant, on tothom pot transmetre sense llicència.

2.5. Nivell de xarxa

La capa de xarxa es troba entre la capa d'accés a la xarxa i la capa de transport. Aquesta capa és la que permet als servidors posar paquets de dades dins d'alguna xarxa i viatjar als seus destins. És a dir, ofereix el servei d'encaminament dels missatges i de traducció de les adreces lògiques en adreces físiques. També efectua el control de la congestió de la xarxa i la reordenació de paquets un cop arribats a la màquina de destí.

2.5.1. OSPF

El protocol OSPF es va desenvolupar com a resposta a la necessitat d'introduir a la comunitat d'Internet un protocol no propietari d'altres funcionalitats per a la família TCP/IP, que suportés xarxes grans i escalables que RIP no pot afrontar. L'IETF el va desenvolupar l'any 1988, però no es va formalitzar fins l'any 1991. La versió més recent és la versió 2 (descrita al RFC2328).

Comparat amb altres protocols interiors, com el RIP, l'OSPF presenta els següents avantatges:

- No té limitacions pel que fa a les dimensions de la xarxa on s'utilitzin (el protocol RIP està limitat a un màxim de 15 salts).
- Permet utilitzar màscares de longitud variable (VLSM, Variable Length Subnet Masks) i per tant, admet la definició de subxarxes. Cal dir que el protocol RIPv2 ja contempla aquesta funcionalitat.
- Les taules d'encaminament només es transmeten a l'inici (quan s'arranquen els encaminadors), quan es produeix algun canvi i cada mitja hora, per assegurar la correcta integritat de la informació. Per tant, no consumeix tant ample de banda com el protocol RIP on la transmissió de la informació d'encaminament és periòdica (cada 30 segons per defecte).
- Presenta una millor convergència que RIP, ja que els canvis es notifiquen a tots els nodes de la xarxa mitjançant inundació i, en paral·lel, es calculen les taules. En el cas de RIP la convergència és lenta a causa de les temporitzacions i perquè les modificacions no es realitzen immediatament.
- Utilitza un valor de cost com a mètrica per seleccionar les rutes. El protocol RIP pren decisions d'encaminament en funció del nombre de salts (vector distància), independentment de la càrrega de la xarxa, de l'ample de banda o del retard. En canvi, l'OSPF utilitza un valor de cost configurable (vector distància).
- Permet la divisió de la xarxa en àrees amb l'objectiu de reduir la transmissió d'informació d'encaminament a tota la xarxa. La separació en àrees implica la necessitat de definir encaminament jeràrquic.

L'OSPF és un protocol basat en un algorisme d'estat d'enllaç. L'estat d'enllaç es transmet mitjançant LSAs (Link State Advertiments) que constitueixen una descripció completa de tota la xarxa i que els encaminadors s'intercanvien amb l'objectiu de confeccionar una base de dades global de la xarxa. Amb aquesta base de dades obtenen un mapa complert de la xarxa i sobre ell calculen les rutes òptimes cap a cada destí utilitzant l'algorisme de Dijkstra. Existeixen diferents tipus de LSAs depenent de la part de la xarxa que es descriu. En xarxes grans, les dimensions de la base de dades podrien representar un problema a l'hora de realitzar els càlculs i a l'hora de difondre la informació a la

resta de nodes. Aquest problema es redueix gràcies a la introducció del concepte d'àrea OSPF. La creació d'àrees redueix la mida d'aquesta bases de dades. Tots els encaminadors que pertanyen a una àrea mantenen una mateixa base de dades i només exporten fora de l'àrea un resum de la mateixa.

2.5.2. BGP

El Border Gateway Protocol (BGPv4) és un protocol d'encaminament entre sistemes autònoms. Es va definir amb l'objectiu de resoldre determinats problemes i limitacions del seu precursor, l'Exterior Gateway Protocol (EGP).

La funció principal d'un sistema que executa BGP és intercanviar informació amb altres sistemes BGP per tal de conèixer el camí cap a diferents xarxes. Aquesta informació inclou una llista amb els diferents sistemes autònoms (ASs) que travessen els missatges. Aquesta llista serveix per diferenciar un graf de la connectivitat entre diferents ASs que permet evitar bucles i establir certes polítiques d'encaminament a nivell d'AS.

Segons la definició clàssica, un sistema autònom és un conjunt d'encaminadors sota una administració tècnica única, que utilitzen un protocol d'encaminament interior (RIP i OSPF per exemple) amb una mètrica comuna per encaminar els paquets dins de l'AS, i un protocol d'encaminament exterior per encaminar els paquets cap a altres ASs. No obstant, a la pràctica, és comú que un únic AS utilitzi més d'un protocol d'encaminament interior i, sovint, més d'una mètrica dins de l'AS. Per això, el terme AS es pot utilitzar per referir-se a un sistema que, malgrat utilitzar diferents IGPs i vàries mètriques es pot percebre des de fora com una unitat d'encaminament coherent. Les xarxes corporatives, com les universitàries o les empresarials, normalment utilitzen un protocol IGP com el RIP o l'OSPF per intercanviar informació d'encaminament entre les seves xarxes, mentre els usuaris connectats a algun ISP, i els propis ISPs entre sí, utilitzen el BGP per intercanviar les rutes d'usuari i d'ISP. Quan el BGP s'utilitza entre ASs diferents, rep el nom d'external BGP (eBGP). Si un proveïdor de servi l'utilitza per intercanviar rutes dins d'un mateix sistema autònom, aleshores s'anomena internal BGP (iBGP).

El BGP és un protocol d'encaminament molt robust i escalable, com ho evidencia el fet que sigui el protocol d'encaminament utilitzat en la Internet. Actualment, les taules d'encaminament BGP tenen més de 10000 rutes. Per aconseguir escalabilitat a aquest nivell, el BGP utilitza diferents paràmetres de ruta, anomenats atributs, per definir polítiques d'encaminament i mantenir un entorn estable.

La versió quarta del protocol està descrita al RFC1771 i proporciona un conjunt de mecanismes per suportar encaminament entre dominis sense classes (o CIDR), això significa que es poden advertir prefixes IP eliminant el concepte tradicional de classes. BGPv4 també introdueix mecanismes que permeten l'agregació de rutes, incloent agregació dels camins dels ASs.

2.6. Propostes de Testbed

Un cop conegut l'escenari i entorn en el qual es vol desenvolupar aquest estudi es realitza un seguit de propostes, el desenvolupament de les quals es detalla durant la resta del document. Aquestes propostes s'han de poder testejar amb el maquinari disponible i tot el software compatible actualment. També és pot pensar en adaptar algun programari en cas de ser necessari, però ambdós distribucions escollides compten amb solucions ja funcionals.

Les propostes s'han de focalitzar en les tecnologies presentades en el capítol anterior i la seva realització ha de simular l'entorn de Guifi.net de la forma més fidedigna possible. També cal dir que la extracció de coneixements, conclusions i qualsevol tipus de recurs software generat (codi, configuració, etc) ha d'ésser d'utilitat per a Guifi.net. En el millor dels casos es pot plantejar una participació activa en la seva implantació, encara que es creu que aquest pot ser un procés llarg i costós.

2.6.1. Encaminament

La proposta pràctica d'aquest estudi sorgeix dels coneixements i experiència adquirits al llarg de diverses assignatures realitzades a l'EPSC, especialment durant el mòdul de Xarxes, Sistemes i Serveis i Laboratori de Xarxes.

Una de les tasques que ha quedat parcialment aturada per part de Guifi.net és el test dels protocols OSPF i BGP per tal d'optimitzar la xarxa i dotar-la de suficient tolerància a fallides, així com balanceig de càrrega. El protocol BGP ha de permetre configurar una Guifi.net on, i per exemple, cada comarca és un sistema autònom, que executa OSPF com a protocol intern i BGP per comunicar-se amb la resta de zones. A la pràctica, aquesta proposta consisteix en continuar amb les proves dels protocols i escenaris amb el hardware disponible per tal d'extreure coneixements, conclusions i configuracions útils.

CAPÍTOL 3. PART PRIMERA: ENCAMINAMENT INTERIOR

3.1. Introducció

L'arquitectura actual d'OSPF de Guifi.net es sustenta en la idea d'utilitzar una única zona, la backbone o troncal. En conseqüència, totes les entitats OSPF de la xarxa estableixen un nivell de veïnatge d'igual a igual, sense existir cap tipus de particionat lògic de la xarxa. Com que tots els encaminadors formen part de la mateixa zona es notifiquen les subxarxes que tenen connectades seguint un esquema tots-contra-tots. Així doncs, cada element té constància aproximadament, de més de 400 subxarxes. Aquesta situació és totalment inviable en quant a escalabilitat de la xarxa. El fet d'utilitzar equips amb capacitats hardware molt limitades porta aquests elements al màxim de les seves capacitats.

Alguns integrants actius de Guifi.net estan testejant configuracions amb aquests protocol, i per tant, existeixen ja alguns coneixements i configuracions útils. Per tal de poder treballar aquesta part de l'estudi focalitzant en els punts més crítics o d'interès per Guifi.net es realitza una reunió inicial per tal d'extreure una llista de punts d'interès. Així doncs, cal fer èmfasi en una sèrie de funcionalitats dels protocols per tal d'adaptar-los a Guifi.net i la resta de programari existent:

- L'objectiu principal i per tant, el resultat que es vol obtenir és reduir el nombre de rutes que ha de tenir en memòria cada node de la xarxa.
- Veure les diferents tipus de zones OSPF i la seva aplicació en la configuració dels nodes de Guifi.net. Actualment, tots els encaminadors obtenen un rol de ASBR de forma automàtica i a priori sembla que aquesta no és la configuració apropiada. De fet, aquest és el primer punt que cal d'aconseguir l'objectiu principal d'aquesta part.
- Automatitzar la configuració dels diferents rols OSPF i possiblement les mètriques de cadascun dels enllaços.
- Un altre punt molt important és definir un criteri per a la definició dels diferents sistemes autònoms (AS) dins de Guifi.net i veure quines implicacions pot tenir en els criteris per definir les diferents zones OSPF.
- Possiblement, també caldrà veure quina és la configuració per tal de poder utilitzar túnels dins de Guifi.net o, si per contra, cal pensar en algun altre tipus de solució.

Per tal de poder estudiar cadascun d'aquests punts de forma individual però no havent de configurar escenaris de forma independent es proposa la realització de cinc escenaris diferents.

Per últim, comentar que els escenaris descrits en el document i els tests realitzats en cadascun d'ells, entre d'altres, es troben degudament documentats i comentats a l'annex segon, "Configuracions OSPF".

3.2. Primer escenari: OSPF i l'àrea troncal.

Aquest primer escenari pretén ésser una còpia de l'actual funcionament de Guifi.net, on tota la xarxa esta formada per una única àrea. La configuració d'aquest escenari servirà per tenir els aparells parcialment configurats per a la resta d'escenaris i per obtenir uns coneixements sòlids sobre el funcionament dels diferents dimonis d'encaminament.

L'àrea 0.0.0.0 ha d'estar present en tota xarxa OSPF, a més que en cas d'existir altres àrees aquestes han d'estar connectades a aquesta. Per una banda, el fet d'utilitzar una única àrea en aquest escenari minimitza els recursos que cadascun dels equips ha d'utilitzar, ja que no hi ha camins redundants i només existeix una única base de dades del protocol, compartida i coherent en tots els equips. Però per contra, el fet d'utilitzar una única àrea implica que totes les rutes han d'estar presents dins la base de dades d'aquesta àrea i per tant, aquesta creixerà tant com ho faci la xarxa. Com ja s'ha comentat, en el cas de Guifi.net, la mida de la base de dades de àrea backbone és de més de 400 subxarxes. A priori, els requeriments de memòria i capacitat computacional van lligats a la grandària d'aquestes bases de dades i per tant, com més gran sigui àrea més recursos es necessitaran en cada un dels equips.

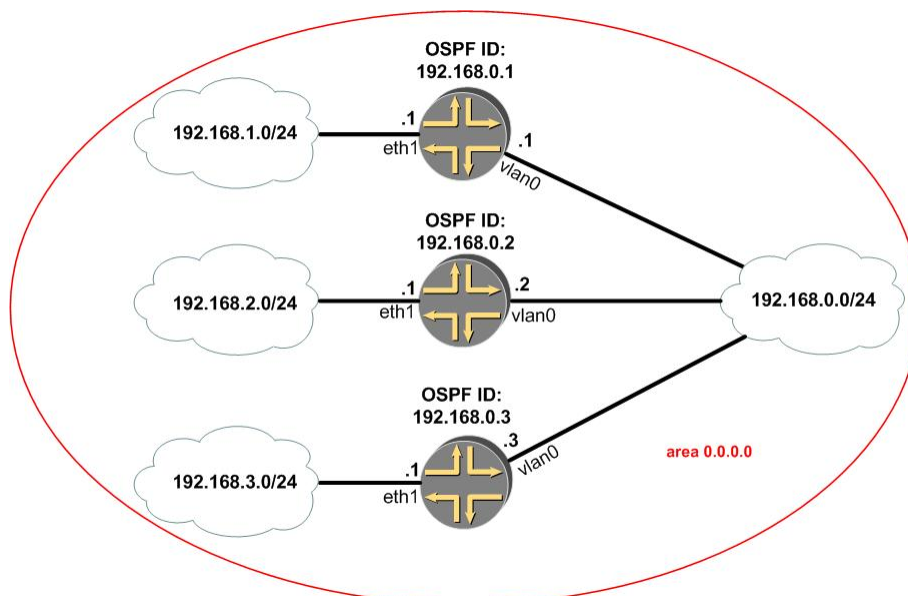


Fig. 3.1 Arquitectura del primer escenari OSPF

La conclusió que s'extreu després d'estudiar aquest escenari ja és coneguda a priori: L'actual arquitectura de Guifi.net no és gens extensible i donat el creixement que té aquesta comunitat s'ha de pensar en trobar solució a aquest punt com a objectiu principal d'aquest estudi.

3.3. Segon escenari: OSPF i particionat àrees.

El segon escenari pretén estudiar els primers objectius comentats a l'inici d'aquest punt. Per a tal, es decideix configurar una xarxa amb més d'una àrea i observar quina és la millor configuració a aplicar en cada cas. És important obtenir una configuració on no tots els equips obtinguin un rol d'encaminador ASBR (Autonomous System Boundary Router) ja que en realitat tot l'escenari es podria comportar com un sistema autònom (AS), o part d'ell, no havent-hi la necessitat d'establir més d'un ASBR.

Els diferents rols que un encaminador pot prendre dins d'una àrea són quatre:

- **Encaminador Intern:** És el responsable del manteniment de la base de dades interna de àrea i per tant, ha d'encaminar totes les subxarxes internes. A més, s'encarrega de traspassar tràfic entre diferents àrees.
- **Encaminador de troncal:** És aquell que té almenys una interfície connectada a àrea troncal i per tant, té coneixement de tota la xarxa, amb o sense sumarització de rutes.
- **Encaminador ABR (Àrea Border Router):** És aquell que té interfícies en més d'una àrea a més de la connexió amb l'àrea troncal. Manté informació sobre aquesta àrea concreta. Si aquesta àrea no té cap altra connexió amb altres àrees s'anomena stub àrea i rep un tractament especial. En aquest últim cas comentat, no és necessari transferir informació sobre la topologia de la xarxa cap a l'interior de àrea, sinó que l'encaminament es realitza a partir de la ruta per defecte.
- **Encaminador ASBR (Autonomous System Boundary Router):** Està localitzat entre una xarxa OSPF i una xarxa no OSPF. Aquest equip és l'encarregat de la distribució de rutes externes (a l'àrea) dins el protocol OSPF i per tant, la resta de la xarxa. Veient les característiques d'una àrea stub de seguida es pot veure que un ASBR no té que estar en aquest tipus àrees.

De la mateixa forma que un encaminador pot prendre un rol dins d'una àrea, OSPF també defineix els diferents tipus àrees, que cal considerar en aquest escenari:

- **Àrea normal:** Una àrea normal és aquella en la que les rutes externes d'aquesta son transmeses a la resta d'àrees a través del respectiu ABR. A menys que en la configuració de àrea s'indiqui el contrari, totes les àrees són d'aquest tipus.

- **Àrea stub:** Una àrea stub és aquella que no té cap més connexió amb altres àrees que no siguin la troncal. Com a conseqüència, l'ABR no retransmetrà les rutes externes de l'àrea troncal cap aquesta sinó que generarà un sumari i una ruta per defecte (sumari LSA de tipus 3) i la transmetrà. Per tant, tot l'encaminament del tràfic sortint d'una àrea stub es realitzarà amb la ruta per defecte, sense necessitat de conèixer la xarxa. A més, una àrea stub no permet tenir cap ASBR per la mateixa naturalesa d'aquests.
- **Àrea totally-stubby:** Un cas particular àrea stub, en la qual només es retransmetrà la ruta per defecte i ni tant sols un sumari des de l'àrea troncal. És habitual configurar aquest tipus d'àrea si dins d'ella no hi ha cap altre encaminador que no sigui l'ABR.
- **Àrea not-so-stubby (NSSA):** Aquestes permeten més flexibilitat a l'hora de retransmetre les rutes externes, ja que es permeten injectar rutes externes des de dins la mateixa àrea però no es permet la injecció de rutes externes des de altres àrees. Però les rutes externes que s'hagin generat dins d'una àrea NSSA si es transmetran a la resta d'àrees a través de l'ABR corresponent, canviant el tipus d'LSA del 7 al 5.

Per a l'assignació d'adreces en aquest escenari es segueix un esquema jeràrquic, ja que el protocol OSPF està pensat per executar-se en entorns d'aquest tipus. Si es vol pensar en utilitzar els mecanismes de sumari de rutes per tal de minimitzar la mida de la taula d'encaminament aquest punt és imprescindible. Per tant, i prescindint totalment de les adreces de tipus A, B, C i D es particiona l'espai 172.16.100.0/16 en subxarxes amb una màscara /30. Donat que tots els enllaços són punt a punt aquesta és la pràctica més recomanable, per tal d'aprofitar al màxim tot l'espai d'adreçament. Per a les xarxes d'usuari es decideix utilitzar un únic rang 192.168.100.0/24 que es subdividit en 3 rangs més petits. Cadascun dels rangs és assignat a les interfícies inal·làmbriques dels equips. Per tant, tot aquest escenari pot ésser vist com un únic punt d'accés que dona servei a una quantitat d'usuaris més elevada del que pot un equip convencional. Normalment aquest tipus de configuracions s'anomenen hot-spot i es troben en punts concrets on la concentració d'usuaris és màxima. A Guifi.net una configuració d'aquest estil pot ésser molt útil en zones urbanitzades o amb una concentració d'usuaris puntual molt elevada, com pot ser el cas d'events socials concrets o celebració de trobades.

La següent figura mostra d'arquitectura IP i la topologia utilitzada en aquest escenari.

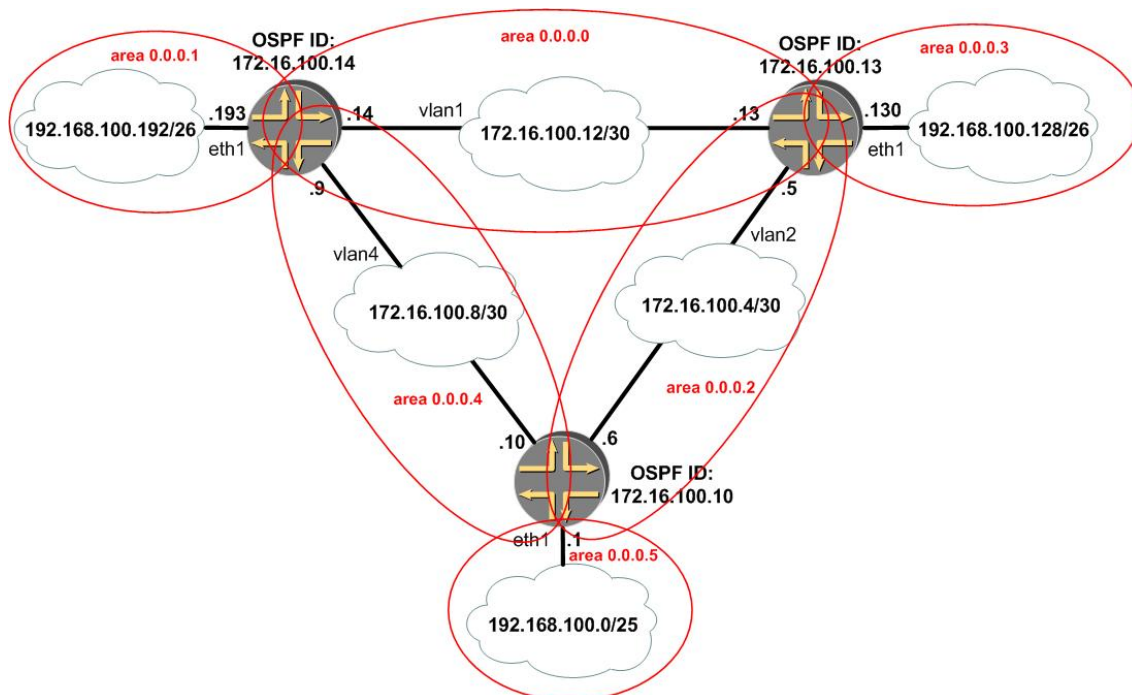


Fig. 3.2 Arquitectura del segon escenari OSPF

Seguint amb la classificació dels diferents tipus d'encaminadors i zones es pot observar que tots els equips d'aquest escenari obtenen un rol d'ABR per a àrea troncal i també d'ABR per a les àrees a les quals estan directament connectats. Totes les àrees a excepció de la 0.0.0.0 estan configurades com a totally-stub, encara que configurar-les com a stub no és un error, però si és ineficient en termes de memòria y arquitectura OSPF.

A l'inici de l'escenari anterior s'ha comentat que tota àrea ha d'estar connectada a la troncal, però no s'ha especificat si aquesta connexió ha de ser física o lògica. OSPF aporta un mecanisme d'enllaços virtuals per a xarxes grans i que s'ha hagut de configurar a l'àrea 0.0.0.5 per enllaçar-la amb l'àrea 0.0.0.0. S'ha comprovat el correcte funcionament d'aquest mecanisme en Quagga i Bird. Amb aquesta funcionalitat, es pot pensar en xarxes OSPF amb moltes àrees.

Aquest escenari permet testejar algunes de les funcionalitats avançades del protocol com és la obtenció de camins alternatius en cas de caiguda d'un enllaç. Per testejar-ho, es desconnecta la vln2 i s'espera que s'esgotin els temporitzadors de veïnatge entre els dos encaminadors d'aquesta VLAN. Un cop l'encaminador 172.16.100.10 i 172.16.100.13 no tenen cap comunicació àrea 0.0.0.2 està trencada i per tant no és accessible. En aquest moment es produeix una redistribució d'informació d'encaminament dins l'àrea 0.0.0.0 i el posterior càlcul del SPT (Short Path Tree) per a cada subxarxa. Aquests càlcul, aplicat a la xarxa 192.168.100.128/25 des de l'encaminador 172.16.100.10 resulta en que el camí més curt és a través d'un equip intermedi, el 172.16.100.14. De fet, si s'observa la base de dades OSPF d'aquest equip

s'observa que no hi ha hagut cap canvi, lògic si s'aprecia que l'enllaç caigut no és contigu a aquest equip.

El valor per defecte dels temporitzadors implicats en el test anterior és de 60 segons i per tant, amb aquesta periodicitat s'envia un paquet de control per cada enllaç (Hello packet) per comprovar el correcte veïnatge OSPF. No és necessari que el valor d'aquests comptadors sigui el mateix en tots els equips, encara que la majoria de fabricants i administradors de xarxa convergeixen en la idea de fixar-lo de forma coherent. Per tant, una conclusió que s'extreu d'aquest escenari és que a Guifi.net cal estudiar amb detall aquest paràmetre. Com més petit es configura el paràmetre més quantitat de tràfic de control es transmetrà en cada àrea. Si la utilització dels enllaços no és pròxima a la seva capacitat no hi ha cap altre motiu que impedeixi disminuir aquest període sense apreciar un empitjorament en el rendiment de la xarxa. Per altra banda, a l'enviar aquests paquets de control amb una freqüència més elevada és té una visió més real de quin és l'estat de la xarxa. Donat que OSPF és un protocol d'estat d'enllaç l'encaminament serà més eficient. Per últim, al minimitzar el temps entre paquets de control, la xarxa s'adonarà més ràpidament de la caiguda dels enllaços, podent reaccionar més ràpidament i restablint la connectivitat per un camí alternatiu.

Per acabar, una de les conclusions més interessants que s'obté d'aquest escenari és que el particionat en àrees és útil, però no cal portar-lo a l'extrem i configurar una àrea per a cada enllaç. A diferència del primer escenari, en aquest cas cada equip obté un rol de ABR (Area Border Router) ja que cada equip és contigu a diferents àrees i per tant, existirà una base de dades diferent per a cada una. Si cada equip pertany a 3 àrees implica que hi ha d'haver 3 bases de dades en cada equip, augmentant la quantitat de memòria i recursos computacionals necessaris. Per tant, per al següent escenari es pensa en configurar com una única àrea els enllaços entre els encaminadors.

3.4. Tercer escenari: OSPF i sumarització de rutes.

El tercer escenari pretén ésser una configuració d'exemple més per a Guifi.net. Aquest pot ésser la reproducció de la xarxa d'un municipi o d'una comarca, ja que s'utilitzen diverses xarxes d'usuari i un particionat en àrees força extensiu.

L'arquitectura és la mateixa que en l'escenari anterior però el particionat lògic de la xarxa és diferent. Als extrems de cada encaminador es configura un mínim d'una àrea per tal de poder testejar la sumarització de rutes i la distribució d'informació en diferents tipus d'àrees. Com que precisament es vol comprovar el funcionament dels sumaris OSPF aquestes xarxes d'usuari han de permetre l'agregació lògica seguint un esquema VLSM (Variable Length Subnetting Mask).

En el cas de l'encaminador 172.16.100.14 es realitza un esquema no del tot jeràrquic per veure si amb una configuració robusta d'OSPF es pot seguir tenint el comportament esperat sense pèrdua d'eficiència. Notar que les dues xarxes

dins de l'àrea 0.0.0.4 podrien agregar-se amb l'adreça 192.168.0.0/X on $X < 24$ però com la 0.0.0.3 està configurada amb la 192.168.0.0/24 inicialment aquesta simplificació no és possible.

La següent figura mostra amb detall el particionat lògic de la xarxa i l'assignació d'adreces a cada una de les xarxes existents. Notar que aquest escenari es pot complicar augmentant el nombre de xarxes als extrems, però la quantitat escollida és suficient per comprovar el correcte funcionament dels mecanismes.

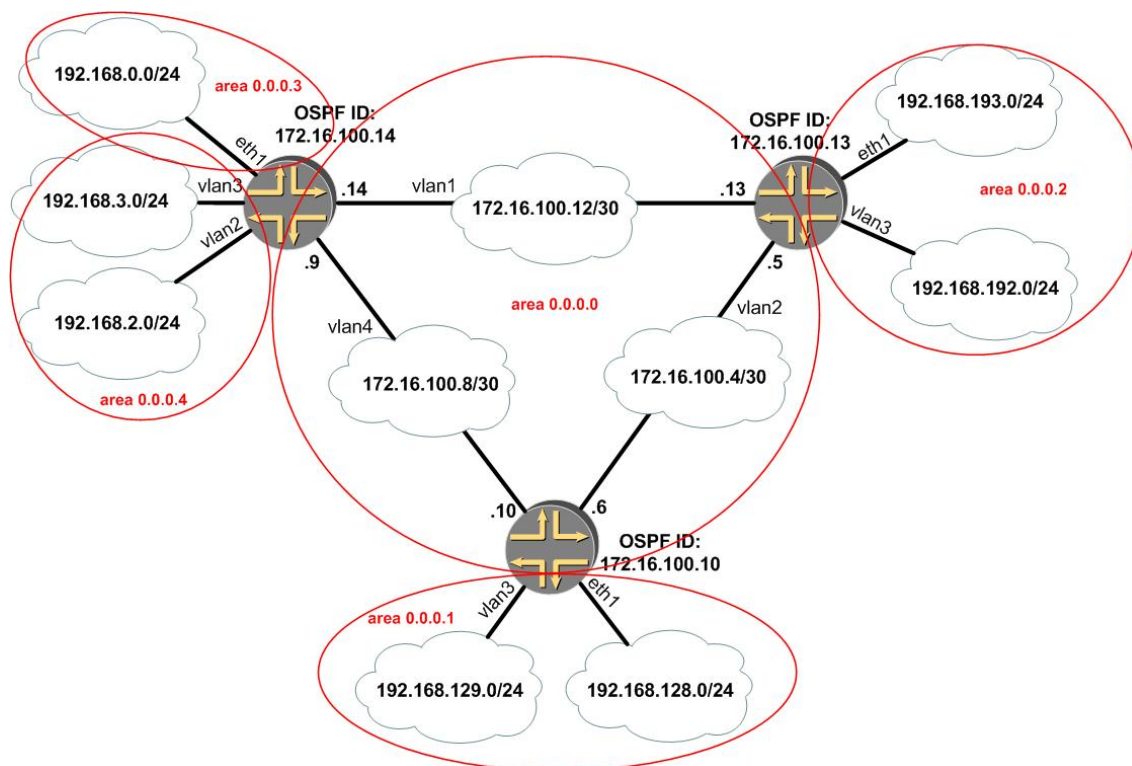


Fig. 3.3 Arquitectura del tercer escenari OSPF.

La sumarització és fàcil de dissenyar en el cas dels encaminadors 172.16.100.13 i 172.16.100.10. Per exemple, i per al primer cas, s'ha configurat el sumari 192.168.192.0/18 a l'àrea 0.0.0.2 en comptes de configurar-la amb 192.168.192.0/24 i 192.168.193.0/24. Els motius són fàcils d'explicar. La primera adreça inclou les dues segones i per tant, la configuració és vàlida. Per altra banda, s'ha disminuït a la meitat la quantitat de rutes que l'àrea 0.0.0.2 injectarà a l'àrea 0.0.0.0. Fins i tot, i si es segueix un pla d'adreçament jeràrquic, amb una única ruta es podrà encaminar fins a 64 subxarxes /24, sempre que aquestes siguin accessibles des de l'àrea 0.0.0.0 a partir de l'àrea 0.0.0.2. Es segueix el mateix procediment en l'encaminador 172.16.100.10 per poder sumaritzar l'àrea amb una única adreça, la 192.168.128.0/18.

Per últim, comentar la configuració que s'ha realitzat en l'encaminador 172.16.100.14 ja que és diferent a les realitzades fins al moment. Dins la configuració de àrea 0.0.0.4 s'ha inhibit la propagació de les dues rutes /24 que la conformen i a àrea 0.0.0.3 s'ha configurat l'equip per a que en comptes

d'anunciar la xarxa 192.168.0.0/24 ho faci amb una màscara /18, com en la resta de casos. Aquest escenari ha estat possible gràcies a que l'àrea 0.0.0.3 i la 0.0.0.4 comparteixen un ABR. En aquest punt, cal observar que els altres dos encaminadors no fan distinció en les subxarxes internes de l'àrea però també és fàcil veure que no requereixen fer-la, donat que el camí és el mateix. Amb aquest procediment s'ha reduït la taula d'encaminament OSPF de cada equip de 10 a 4 subxarxes. També cal dir, que el nombre de rutes existents a la troncal no variarà ni que s'assigni tot l'espai d'adreçament 192.168.0.0/16 a l'escenari, sempre que es faci seguint un model jeràrquic.

Igual que en l'escenari anterior es comprova el correcte funcionament de la tolerància a fallides, per mitjà de diferents tests abaixant enllaços i veient si el tràfic és reencaminat correctament. La convergència es realitza més ràpidament donat que els tres encaminadors pertanyen a la mateixa àrea i que la quantitat d'informació a calcular i transmetre és menor.

Finalment, comentar que el mecanisme de sumarització de rutes permet reduir la mida de la taula d'encaminament dels enrutadors. Aquesta reducció està estrictament restringida a la topologia IP de la xarxa i per tant, un escenari on no es diferenciï entre adreces de troncal i d'usuari, i es realitzi una assignació totalment jeràrquica serà el quin en podrà treure majors avantatges.

3.5. Quart escenari: OSPF i balanceig de càrrega

El quart escenari és una extensió de l'anterior, al qual s'ha configurat un enllaç redundat entre dos dels encaminadors. L'objectiu inicial és testejar-lo com a backup del primer però també es vol comprovar el correcte funcionament del balanceig de càrrega mantenint els dos enllaços actius.

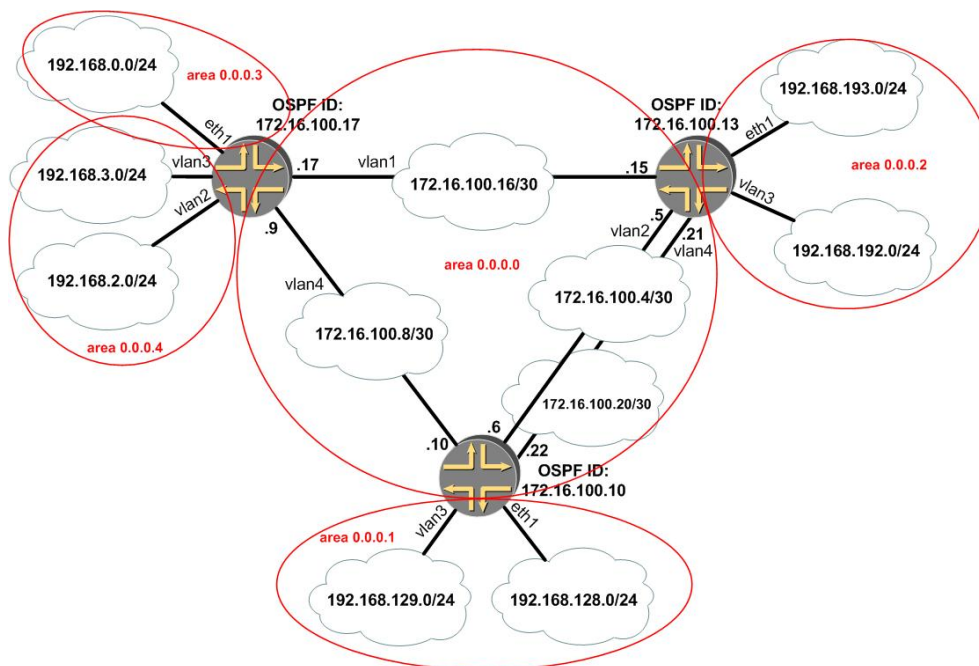


Fig. 3.4 Arquitectura del quart escenari OSPF.

Tal i com es pot observar en la figura 3.5, s'ha configurat un escenari addicional entre un dels parells d'encaminadors. L'assignació d'adreces a aquest enllaç s'ha realitzat seguint la mateixa política que fins al moment, per tal d'impactar el mínim possible en la mida de les bases de dades i taules d'encaminament.

Durant la realització d'aquest escenari s'ha estudiat el comportament de l'OSPF per diferents mètriques d'enllaços:

- D'una banda, configurant les interfícies pertanyents a la xarxa 172.16.100.20/30 amb una mètrica de valor 15, més elevada que la resta, s'aconsegueix que el camí redundat resti inactiu. Si el camí actiu (mètrica 20) no és accessible s'utilitzarà el de backup (mètrica 25) en comptes d'escollir la tercera ruta possible (realitzant un salt intermedi, mètrica 30) que es generarà en cas que també l'enllaç de backup tampoc estigui disponible.
- Per contra, si totes les mètriques de l'escenari són iguals existeixen dos camins amb la mateixa mètrica entre dos dels enrutadors. Amb això, cap enllaç realitza funció de backup, sinó que els dos enllaços resten actius. En termes OSPF aquest mecanisme s'anomena balanceig de càrrega. Tot i que actualment moltes de les implementacions de TCP ja suporten paquets fora d'ordre, en el cas d'UDP els paquets desordenats impliquen una pèrdua de qualitat notòria. Per tant, no es considera un mecanisme molt útil si els canals utilitzats són ràdio, amb característiques molt canviants i que possiblement generaran una gran quantitat de retransmissions i desordenaments.

Com a conclusió a la realització d'aquest escenari es pot veure que la capacitat d'assignar diferents mètriques als enllaços és el mecanisme que permet una certa enginyeria de tràfic, aplicat a tot el flux IP, i sense diferenciar usuaris o connexions [1]. Per tant, es pot considerar un mecanisme molt interessant i a tenir el compte a l'hora de dissenyar cada àrea i també la totalitat de la xarxa OSPF.

3.6. Cinquè escenari: OSPF extensiu

El cinquè i últim escenari parteix de la idea de construir un escenari on almenys un dels encaminadors estigui completament dins d'una àrea que no sigui la troncal. D'aquesta manera es podran estudiar amb més detall els mecanismes de sumaris de rutes i la propagació de la ruta per defecte, mecanisme disponible en OSPF. Tanmateix, també es vol comprovar que una àrea stub només necessita una ruta per funcionar. A més, al llarg de tots els escenaris s'han testejat tot tipus d'àrees a excepció de les NSSA (Not-So-Stubby Àrea) i es considera aquest escenari com últim per poder testejar les seves funcionalitats.

Com ja s'ha explicat, una àrea NSSA realitza una conversió entre diferents tipus de LSA de forma que és capaç d'introduir dins l'OSPF rutes que inicialment pertanyen a altres protocols (RIP o BGP). Aquesta funcionalitat no s'ha comprovat de forma exhaustiva, donat que no es disposa d'un quart

encaminador per generar rutes externes a la xarxa de l'escenari cinquè. Per a tal, s'ha desconfigurat una de les xarxes de l'enrutador per afegir-ne una altra que inicialment no està contemplada dins la configuració OSPF. Com aquesta xarxa és externa, la ruta que genera també es considera com externa i inicialment no retransmesa a la resta de la xarxa. Després de configurar l'àrea com a NSSA aquesta ruta es distribuïda per tota la xarxa OSPF.

La següent funcionalitat que s'ha volgut testejar és la creació i distribució de la ruta per defecte (0.0.0.0/0) dins la xarxa. Per a tal objectiu, s'ha configurat l'enrutador de l'àrea NSSA com a generador d'aquesta ruta. Després de reiniciar el procés OSPF aquesta ruta s'ha retransmès cap a l'encaminador veí, i d'aquest a l'enrutador extrem, de l'àrea stub.

Seguint amb l'últim objectiu, aquesta ruta per defecte també ha estat retransmesa dins l'àrea 0.0.0.1, configurada com a stub. Arribat a aquest punt, la resta de rutes són prescindibles, donat que qualsevol xarxa no pertanyent a aquesta àrea és accessible únicament a través de l'ABR. Per comprovar el correcte funcionament d'aquest mecanisme és necessari comprovar que les rutes de l'àrea 0.0.0.2 no apareixen dins l'àrea 0.0.0.1 però aquestes continuen essent accessibles gràcies a la ruta per defecte. Així doncs, s'ha comprovat que es pot reduir dràsticament la quantitat de rutes dins d'una àrea si aquesta es configura com a stub i a més, s'hi propaga una ruta per defecte.

La següent figura mostra l'arquitectura configurada per a l'últim escenari d'OSPF:

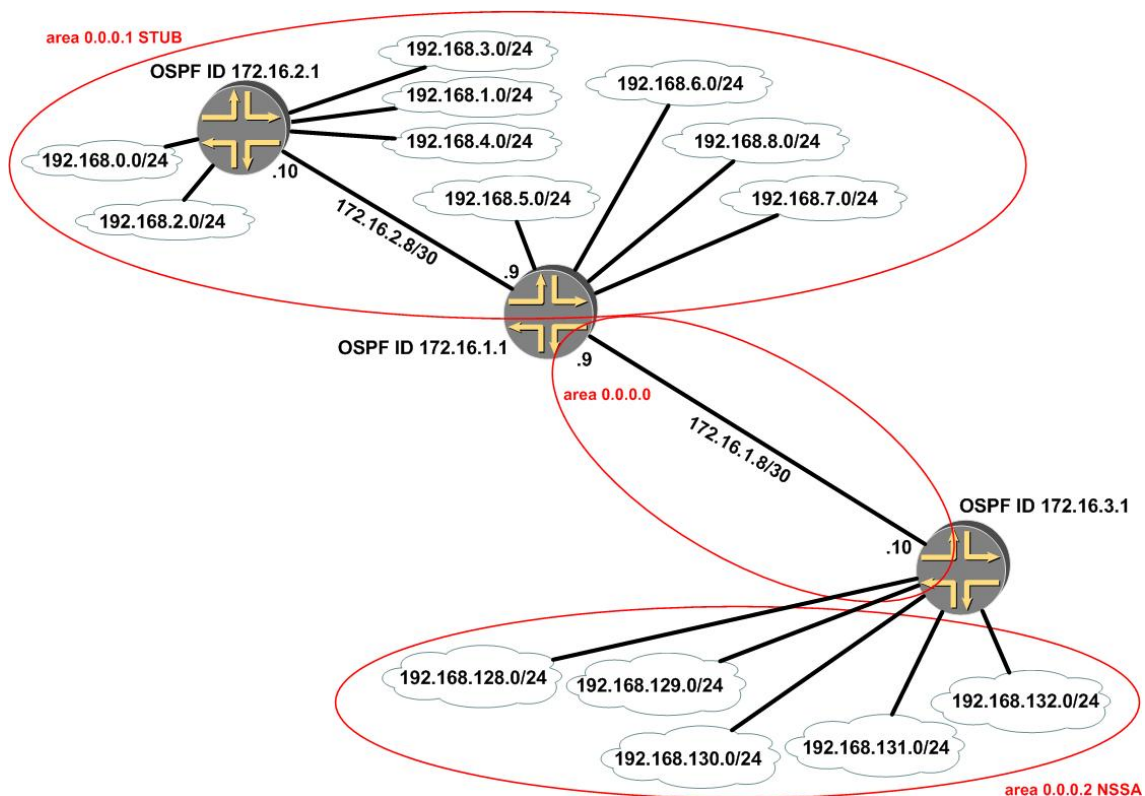


Fig. 3.5 Cinquè escenari OSPF.

Com en els escenaris anteriors, també s'ha configurat la sumarització de rutes. Un cop més, s'ha pogut comprovar l'eficiència d'aquest mecanisme donat que l'adreçament s'ha assignat de forma jeràrquica. Així doncs, la quantitat de rutes diferents a les generades per xarxes directament connectades ha estat de 5, 7 i 2 respectivament i d'esquerre a dreta seguint el diagrama. Cal dir, que en tot l'escenari s'han adreçat (no assignat) fins a 255 subxarxes amb una quantitat de rutes molt menor.

Per mitjà de la configuració d'àrees Stub i NSSA es pot comprovar que aquestes són molt interessants, ja que la configuració d'aquests rols de forma adequada impacta directament en el bon funcionament de l'OSPF fins la xarxa i la optimització dels recursos utilitzats. Així doncs, es recomana:

- Configurar àrees Stub a l'extrem de la xarxa OSPF. La configuració d'aquest rol a les àrees més externes permet reduir dràsticament la quantitat de rutes introduïdes dins l'àrea. Els encaminadors interiors només coneixen la topologia de la pròpia àrea i per tant, cal acotar la seva mida. L'encaminador ABR de àrea serà l'encarregat d'injectar-hi la ruta per defecte.
- Configurar àrees NSSA per poder integrar l'encaminament interior i l'exterior. Partint de la base que tot AS, que executarà OSPF com a protocol interior, tindrà connexió amb altres ASs caldrà injectar les rutes externes del BGP dins de cada xarxa OSPF. Aquest mecanisme permet que no tots els encaminadors de la xarxa OSPF hagin d'executar el protocol BGP, sinó que es pot acotar la quantitat d'encaminadors que l'utilitzin dins d'una o diverses àrees NSSA.

3.7. Consideracions de rendiment

Des del punt de vista tècnic, els protocols d'encaminament permeten un consum de recursos proporcional al creixement de la xarxa. Com més petita sigui aquesta pendent, es considera que el rendiment d'aquesta amb una arquitectura donada és màxim. A la pràctica, els tres recursos a tenir en compte són l'ús de memòria, el consum de CPU i la quantitat de tràfic injectat a la xarxa.

Els protocols utilitzen la memòria per emmagatzemar les rutes i la topologia de xarxa. L'ús del mecanisme de sumaris permet acotar el consum de memòria en tots els protocols d'encaminament i per tant, també en l'OSPF. Donat que la memòria física que ocupa una ruta dins l'encaminador és la mateixa (uns 4k bytes), independentment de la màscara o tipus d'adreça el millor mecanisme per limitar el creixement del consum és acotar la quantitat de rutes. Com ja s'ha demostrat, el millor i més efectiu mecanisme existent per a tal és la partició en àrees i la configuració dels sumaris entre elles.

Per altra banda, el consum de CPU sí és dependent del protocol. La quantitat de càlculs que ha de realitzar cada encaminador OSPF en un entorn amb n

LSA (Link State Advertisement) és proporcional a $n \times \log(n)$. Per tant, es pot afirmar que en una xarxa OSPF el cost computacional és dependent del nombre d'encaminadors i del nombre d'enllaços. A la pràctica s'ha comprovat que el factor més influent és el nombre d'encaminadors encara que també ho són la bidireccionalitat d'aquests enllaços i la taxa d'error entre d'altres. Per tant, la primera expressió donada és aproximativa però vàlida. Alguns protocols utilitzen cicles lliures del processador per comparar noves rutes amb les existents o per calcular noves rutes abans d'un canvi en la topologia. Per als protocols d'estat d'enllaç com és l'OSPF, mantenir les àrees petites i utilitzar sumaris redueix l'efecte d'un canvi en la topologia ja que el nombre de rutes que s'han de tornar a calcular és mínim. Durant aquest estudi, s'ha observat que el protocol OSPF executat en els WRT54GL sota les opcions Quagga/Bird consumeix fins al 40% del temps de processador quan hi ha una retransmissió d'informació i l'algorisme SPT s'executa per recalculat el millor camí cap a cada xarxa. La implementació actual de l'OSPF sota Quagga té deficiències d'optimització del càlcul del SPT [2], encara que s'han desenvolupat algunes extensions que igualen el seu rendiment al de reconeguts equips comercials. Una altra alternativa és utilitzar el comptador de backoff exponencial per a OSPF, encara que aquesta opció implica un compromís. L'ús d'un backoff segons aquesta proposta [3] modifica el comportament de l'OSPF per millorar el seu comportament en la primer i fins i tot segon event dins la xarxa, però aporta una convergència i rendiments baixos quan els events es produeixen de forma periòdica i seguida.

L'ús de l'ample de banda també és dependent del protocol i per al cas d'OSPF cal considerar dos punts de disseny que implicaran en un consum d'ample de banda:

- Les actualitzacions periòdiques s'enviaran de forma regular. Quan les característiques del canal son canviants es recomana un interval baix, de 20 – 25 segons per contra els 30 o 60 d'algunes implementacions. Les actualitzacions completes només es realitzen quan hi ha una alteració en la topologia de la xarxa i per tant, com més estable sigui la xarxa menys ample de banda es consumirà.
- Totes les actualitzacions s'envien a direccions multicast concretes i per tant, si la xarxa ho suporta l'eficiència també augmentarà. Donat que Guifi.net es sustenta en un entorn WLAN tampoc cal incidir de forma especial en aquest tema, ja que el medi es compartit per tots els participants en la xarxa.

Per millorar l'ús que l'OSPF fa de l'ample de banda també existeixen algunes recomanacions, les quals es centren en la implementació d'un OSPF adaptatiu. Aquesta modificació en basa en la idea d'actualitzar el cost de cada enllaç en funció de l'ample de banda efectiu estimat en cada interfície activa de l'encaminador. També s'ha demostrat, que aquesta proposta augmenta la capacitat de la xarxa fins en un 25%, dependent de l'escenari [4].

3.8. Conclusions

Aquest punt pretén realitzar un recull dels punts importants a l'hora de realitzar un disseny d'OSPF robust aplicat a Guifi.net, però sense voler donar una solució o proposta final donat que encara no s'ha estudiat el protocol BGP.

OSPF és un protocol escalable que apareix com a solució al creixement de les xarxes IP. El particionat de la xarxa permet reduir parcialment la complexitat de la xarxa quan aquesta creix. Els dos punts forts d'OSPF que s'han pogut comprovar són la ràpida i eficient convergència i la robustesa de la base de dades de LSAs.

Tanmateix, també s'ha comprovat que l'escalabilitat d'una xarxa OSPF parteix d'utilitzar un esquema d'adreçament jeràrquic en la major mesura possible, ja que és la solució per minimitzar la quantitat de rutes. També s'ha comprovat que aquest factor afecta directament al consum de recursos de memòria i computació.

El primer i més important dels punts a considerar a l'hora de dissenyar una xarxa OSPF és la determinació de quins encaminadors i enllaços formaran part de l'àrea troncal. D'aquesta consideració, en resultaran la quantitat de bases de dades existents en la xarxa i el temps de convergència i distribució entre aquestes.

La quantitat d'encaminadors en cada àrea també cal ésser considerada. Com ja s'ha vist, la quantitat de càlculs a realitzar és major com més gran és l'àrea. Per tant, les àrees que mantinguin un comportament estable es poden mantenir de mides significativament grans (no més de 30 equips) mentre que les àrees que tinguin enllaços s'han de mantenir el més acotades possible. Per aquest últim cas, acotar l'àrea a un enllaç problemàtic pot aportar la màxima estabilitat.

La quantitat de veïns de cada encaminador és important ja que OSPF envia tots els LSA a tots els encaminadors de l'àrea i per tant, com major és la quantitat de veïns major computació és necessitarà quan hi hagi un canvi en la topologia a causa de la recepció de múltiples LSA. En general es recomana que un node no ha de tenir més de 40 veïns. En cap cas es recomana una estructura de tots contra tots.

La quantitat d'àrees en la que es veu involucrat un equip afecta els seus requeriments, ja que s'executa l'algorisme d'estat d'enllaç per a cada àrea. Per tant, i en la mesura possible, és recomana que els equips amb més de 5 àrees siguin equips amb capacitats majors que les del WRT54GL o semblants.

La configuració final d'OSPF ha d'ésser robusta i per tant, cal veure l'efecte de modificar els valors dels comptadors i de les mètriques dels diversos enllaços. Una altra pràctica, no utilitzada durant la realització dels escenaris però si és interessant comentar, és la d'utilitzar el mateix adreçament per a usuaris i troncal. La conseqüència és una reducció de les taules d'encaminament, ja que les xarxes troncales ja estan incloses dins la xarxa sumaryada de tota àrea o xarxa OSPF.

Com a experiència de la configuració dels diversos escenaris s'ha pogut comprovar que és important configurar cada àrea com el que segons el disseny OSPF és (stub, nssa, etc..) ja que incideix directament en el rendiment i bon funcionament de l'encaminador a l'hora que minimitza els recursos necessaris per a fer-ho.

Encara que moltes de les implementacions de TCP suporten paquets fora d'ordre. Encara en alguns casos i en tots en el cas d'UDP el desordre de paquets que pot generar un balanceig serà vist com un empitjorament del funcionament de la xarxa i no es recomana el seu ús de forma extensiva. Així doncs, si es volen utilitzar mecanismes per oferir balanceig de càrrega es recomana que aquest es faci a nivells inferiors, concretament a nivell d'enllaç. Sota les implementacions Linux, la capacitat d'agregar dues o més interfícies físiques diferents en una mateixa rep en nom de bonding [5].

CAPÍTOL 4. PART SEGONA: ENCAMINAMENT EXTERIOR

4.1. Introducció

A l'inici d'aquest estudi, cap element de la xarxa Guifi.net utilitza BGP com a protocol per propagar les rutes internes de la pròpia xarxa cap a les altres xarxes o sistemes autònoms. Anàlogament a la primera part d'aquest estudi es realitza una reunió amb integrants de Guifi.net per veure quins són els punts més importants a estudiar. La llista de qüestions interessants a curt termini és la següent:

- Complerts els objectis amb el protocol OSPF també s'haurà d'estudiar la interconnexió entre diferents AS, possiblement amb la configuració d'ASBR (Autonomous System Boundary Router) en certs nodes. Com ja s'ha vist, les rutes del BGP s'injectaran dins l'OSPF per mitjà de les àrees NSSA.
- Sobre el protocol BGP també cal estudiar la possibilitat d'utilitzar els sumaris per a, un cop més, minimitzar la taula d'encaminament de cada node. Cal veure si és més apropiat utilitzar els sumaris automàtics o realitzar una gestió més manual d'aquest paràmetre.
- També es considera interessant poder comprovar els mecanismes per poder realitzar un filtrat de les rutes, utilitzant llistes de control d'accés (access-list), mapes (route-map) i prefixos de ruta (prefix-list). Veure si es pot realitzar una gestió diferenciada entre la entrada i la sortida en una mateixa interfície.

Per ampliar cada escenari el document compta amb un annex associat, on es realitzen comentaris sobre la configuració i algunes de les parts d'aquest apartat.

4.2. Escenari 1: BGP com a protocol exterior (eBGP)

El primer escenari on s'estudia el comportament del protocol BGP en els WRT54GL té com objectiu principal l'obtenció d'una configuració estable i funcional del protocol. Tanmateix, es vol dissenyar un escenari útil, que aportí algun valor a aquest estudi i que serveixi com a punt de partida per als següents escenaris.

Les rutes apreses a través de BGP tenen associades unes propietats, utilitzades per determinar quin és el millor camí entre un origen i un destí quan existeixen múltiples camins. Aquestes propietats s'anomenen atributs de BGP i són els següents:

- **Weight:** Aquest atribut va ésser definit per Cisco i actualment la majoria de fabricants o implementacions el contemplan. El significat d'aquest atribut és local i per tant mai es notifica als veïns. Si un encaminador coneix més d'un camí per arribar a una destinació, escollirà aquell que tingui un weight major.
- **Local Preference:** Aquest atribut s'utilitza per determinar un punt de sortida des de l'AS local. A diferència de l'atribut anterior, aquest sí que es propaga per la xarxa interna; és a dir, dins de l'AS. En cas d'existir més d'un punt de sortida de l'AS, s'utilitza aquest atribut per determinar quin d'ells serà el preferit i per tant, utilitzat de forma regular.
- **Multi-Exit Discriminator:** Aquest atribut, s'anomena MED o atribut de mètrica. El valor d'aquest atribut conté una recomanació sobre quin ha de ser el punt d'interconnexió preferit per l'AS local. Tanmateix, la selecció de ruta a l'AS remot es pot basar en altres atributs de BGP i per aquest motiu el valor és una recomanació.
- **Origin:** Aquest atribut indica com la ruta ha estat apresada pel BGP, ja que existeixen diferents mecanismes. Els valors que pot prendre són 3: iBGP, quan la ruta ha estat apresada per mitjà del BGP i correspon a una xarxa interna a l'AS. El segon valor possible, eBGP, denota que la ruta ha estat apresada també pel BGP però és externa a l'AS local. Per últim, el valor "incomplete" denota que l'origen no és conegut o la ruta ha estat apresada per altres mecanismes no BGP.
- **AS-Path:** Quan la notificació d'una ruta o xarxa travessa un AS per complet, l'identificador d'aquest s'insereix al final d'aquest atribut. D'aquesta forma, es construeix una llista dels diferents AS que ha de travessar el tràfic entre origen i destí.
- **Next-Hop:** En el cas de BGP entre diferents ASs (eBGP) aquest atribut conté l'adreça IP del pròxim salt. En canvi, en el cas de iBGP s'insereix l'adreça IP del pròxim salt fora de l'AS.
- **Comunitat:** Aquest atribut permet agrupar diferents destinacions per mitjà d'un mateix identificador. A fins pràctics, aquest mecanisme permet assignar unes mateixes polítiques d'acceptació, preferència o redistribució de forma conjunta i aunada. Els valors utilitzats són reduïts, en funció de l'abast que tindrà la propagació d'aquesta ruta: El primer valor és "no-export" i en aquest cas, la ruta no s'advertirà als veïns eBGP i per tant, només serà coneguda dins l'AS. El segon valor és "no-advertise" i s'utilitza per indicar a la entitat BGP que no anunciï aquesta ruta a cap veí, ja sigui extern o intern. Per últim, el valor "Internet" denota que aquesta ruta s'ha de propagar de forma global per la xarxa.

Un cop coneguts tots els atributs, estudiats al llarg de tots els escenaris, que té un ruta també és important veure com és seleccionen. Donat que un encaminador BGP pot rebre múltiples notificacions per la mateixa ruta de

diferents fonts, el mecanisme de selecció de ruta no és tant directe o trivial com en el cas de OSPF:

- Si el camí conté un next-hop inaccessible es descarta l'actualització.
- Es prefereix el camí amb l'atribut weight major.
- Si existeixen dos camins amb el mateix weight s'escull el que té una local-preference major.
- Si les prioritats locals son iguals s'escull el camí originat pel BGP local de cada encaminador.
- Si no existeix ruta generada de forma local, s'escull el camí amb un AS-path menor o més curt.
- Si hi ha camins amb la mateixa quantitat de ASs s'escull el camí amb l'atribut origin menor. En aquest cas, iBGP és el menor valor possible, seguit de eBGP i "incomplete".
- Si l'atribut origin també es repeteix, s'escull el camí amb un menor MED.
- Tanmateix, si existeixen camins amb valors de MED iguals s'escull el camí a través del veí més pròxim.
- Finalment, es diferencia per adreça IP i s'escull el camí amb un BGP router ID menor.

Coneguts aquests principis de funcionament, es dissenya un escenari on s'executarà BGP com a protocol encarregat d'anunciar les rutes de 2 ASs diferents entre ells. Tanmateix, aquest escenari encara manté una part important d'OSPF, ja que es considera també important estudiar com transferir les rutes internes de cada xarxa que executa OSPF com a IGP al BGP. La següent figura mostra l'arquitectura OSPF i BGP així com la topologia de l'escenari:

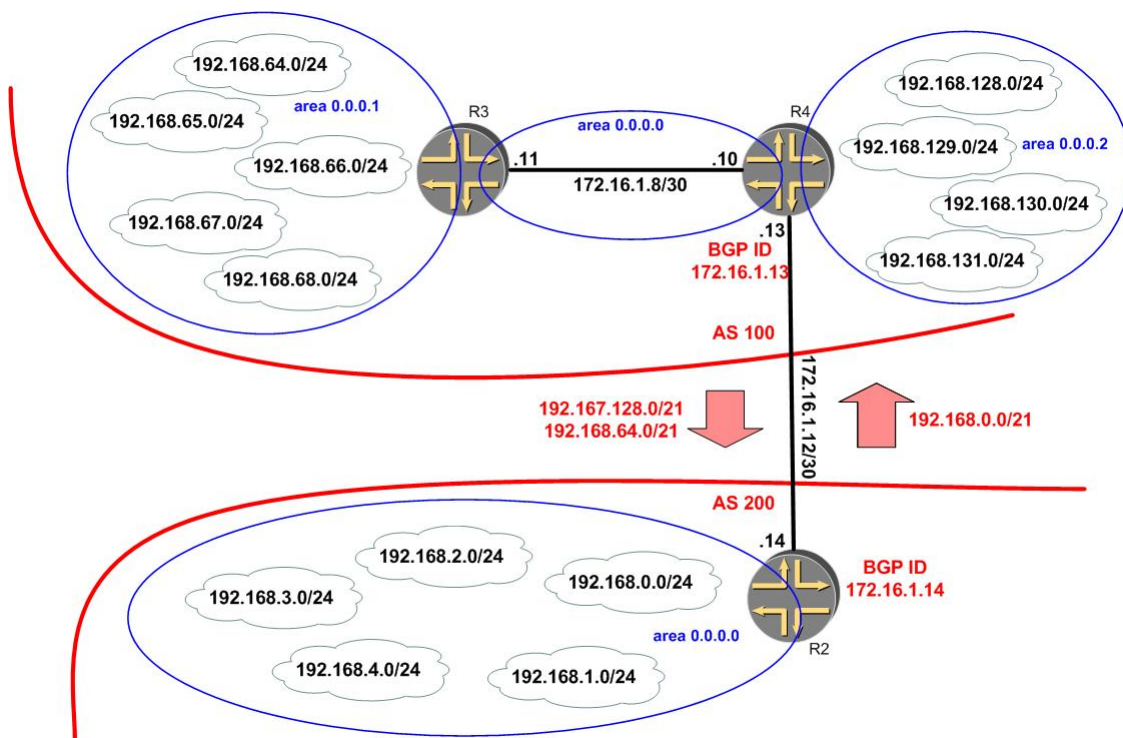


Fig. 4.1 Primer escenari de BGP

Com es pot observar, ambdós xarxes tenen una arquitectura d'OSPF força senzilla, encara que suficient per poder utilitzar sumarització dins del propi AS. La base de dades 0.0.0.0 dins de l'AS 200 és única donat que només hi ha un encaminador i, per aquest motiu, només existeix una única àrea. Per altra banda, dins l'AS 100 si s'han configurat 3 àrees diferents. Donat que a aquest punt de l'estudi ja es coneixen els mecanismes de sumari, s'ha configurat dins del mateix AS i independentment de la sumarització que posteriorment aportarà BGP.

En quant al BGP, s'han planificat 2 AS diferents, enllaçats per un únic camí possible. Notar que en el cas de l'AS 100 només un dels dos encaminadors executen BGP i per tant, s'hauran d'introduir les rutes apreses per aquest protocol dins l'OSPF si es vol tenir connectivitat entre totes les xarxes de l'escenari. Realitzant un anàlisi més extensiu de la es pot veure que les dues xarxes que l'AS 100 anunciarà al veí de l'AS 200 seran diferents en quant als seus atributs, concretament el camp Origin. Com a conseqüència que l'encaminador R3 no executa BGP, la ruta sumaritzada de les seves xarxes estarà apresada com a "incomplete", mentre que les rutes associades a l'àrea 0.0.0.0 i 0.0.0.2, també sumaritzades, estaran presents com a iBGP en la taula d'encaminament BGP del mateix AS.

També es pot notar, que la xarxa 172.16.1.12/30 és imprescindible dins la base de dades BGP de cada un dels nodes, encara que no necessiten intercanviar-la. Aquesta simplificació és possible donat que els identificadors BGP que s'han utilitzat son els mateixos que les IPs que permeten la connectivitat entre els 2 ASs i no cal cap ruta addicional. Per tant, cadascun dels dos encaminadors que executen BGP ja tenen constància d'aquesta xarxa sense necessitat d'advertir-la. Aquesta pràctica només és habitual en topologies molt senzilles i per tant, aplicables a aquest escenari. En posteriors escenaris es mostra una altra pràctica, àmpliament utilitzada, que utilitza com a identificador BGP una IP associada a una interfície de retorn o local.

S'ha comentat que un punt a estudiar dins d'aquest escenari és la integració de l'OSPF i el BGP com a protocols d'encaminament interior i exterior, respectivament. Donat que l'OSPF només realitza decisions basant-se en la taula d'encaminament pròpia d'aquest protocol és necessari introduir aquestes rutes externes dins de la base de dades interna [6]. Per exemple, per al cas de la ruta que l'AS 200 anuncia a l'AS 100, l'encaminador R4 s'encarrega de propagar-la cap a la base de dades de l'àrea 0.0.0.0 i conseqüentment, cap a l'àrea 0.0.0.1. Notar però, que qualsevol altre encaminador situat dins l'àrea 0.0.0.1 ja no tindrà constància d'aquesta ruta ja que el tràfic s'encaminarà cap a l'ABR, encarregat del tràfic dirigit fora de l'àrea.

Després de la realització i estabilització d'aquest escenari s'ha estudiat el comportament de BGP com a protocol external (eBGP) i s'ha comprovat que tot i donada la seva complexitat també es poden realitzar configuracions senzilles i que permeten un disseny escalable en quant al nombre d'encaminadors que executen el BGP i el nombre de rutes que es transferiran entre els diferents ASs. A diferència d'OSPF, el protocol BGP permet trencar amb l'assignació d'adreces IP jeràrquica. El creixement de les taules és igual al de xarxes

anunciades però alhora independentment a la xarxa concreta que s’anuncii en cada una, permetent l’assignació no jeràrquica a cada AS diferent. Per a una bona escalabilitat el protocol BGP pot anunciar les xarxes sumaritzades de les diferents àrees OSPF internes o una xarxa resultat de la sumarització de les diferents xarxes ja sumaritzades.

També s’ha comprovat el funcionament del mecanisme de llistes per poder filtrar les xarxes que cada AS anuncia. El procediment és semblant a la realització d’un tallafocs (firewall) però no aplicat al flux de paquets IP que travessen la xarxa sinó a les rutes que s’anuncien els protocols d’encaminament. Aquest mecanisme pot resultar d’especial interès a l’hora de realitzar una planificació de quines xarxes seran accessibles des d’unes altres.

4.3. Escenari 2: BGP com a protocol interior (iBGP) i exterior

El segon escenari es dissenya com una extensió de l’anterior, però augmentant en complexitat BGP. Aquesta aparentment senzilla modificació permet realitzar una enginyeria complexa amb els atributs MED, Local Preference o altres per instal·lar les xarxes dins la taula d’encaminament. En aquest moment l’estudi se centra en la interacció entre el BGP com a protocol interior i exterior.

La següent figura mostra l’arquitectura i configuració de l’escenari inicial sobre el qual s’han realitzat les posteriors proves. Notar que com existeixen àrees amb adreçament correlatiu apareixen diversos nivells de sumarització possibles (una xarxa /23 per àrea o una xarxa /22 cada dos àrees). A més, es disposa de gran varietat de xarxes a anunciar pel BGP.

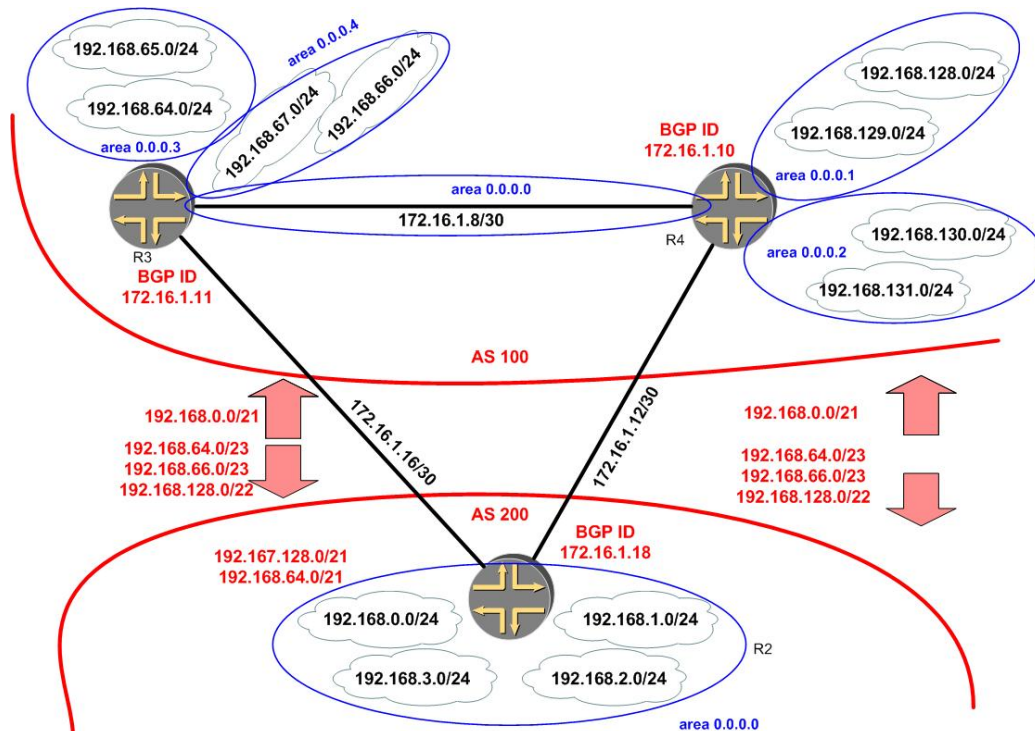


Fig. 4.2 Segon escenari de BGP

Inicialment s'han configurat punts d'entrada diferents per a cada un dels rangs anunciats, donat que aquesta és la configuració òptima si els enllaços són de capacitat baixa i no es vol limitar al BGP a que instal·li un únic punt d'entrada a cada AS. Tanmateix, també s'ha comprovat el comportament habitual del BGP i com la configuració dels atributs permet modificar-ne el comportament. A continuació es detallen un parell de configuracions addicionals i com afecten al camí que segueix el tràfic.

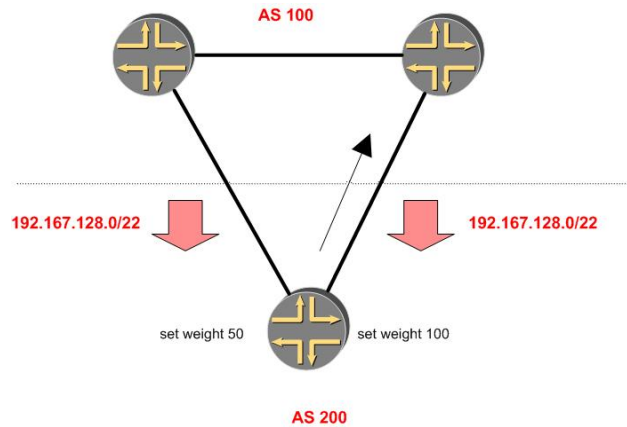


Fig. 4.3 Comportament de l'atribut weight

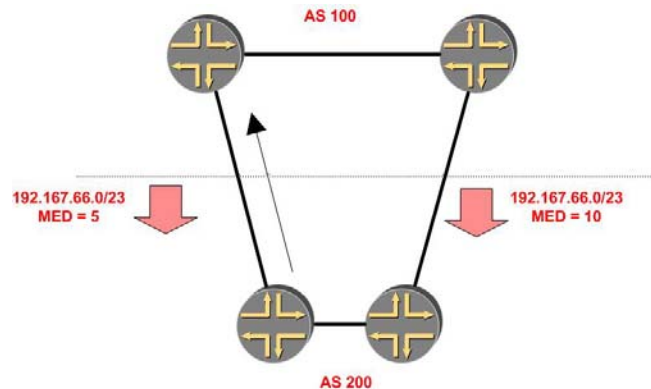


Fig. 4.4 Comportament de l'atribut MED

Per mitjà de les proves realitzades es pot afirmar que iBGP i eBGP són el mateix protocol, però amb diferents regles i finalitats. La diferència principal recau en que l'iBGP no anuncia les rutes externes a la resta de veïns. El motiu és evident si es pensa que no existeix cap mecanisme per a la detecció de bucles en iBGP. Notar que en eBGP totes les rutes s'anuncien a tots els veïns i per tant existeix el risc de la creació d'aquests bucles.

Una altra conclusió extreta és que la multitud d'atributs del BGP permeten assignar pesos i valors de forma que la selecció de la millor ruta és coneguda a priori. Per mitjà dels atributs es pot modificar el comportament d'un AS amb els seus veïns, així com el punt d'entrada i sortida del tràfic per a cada xarxa anunciada. Aquest mecanisme és molt interessant a l'hora de dissenyar una xarxa extensa ja que permet una certa enginyeria de tràfic. S'ha comprovat que

aquests valors s'han d'assignar amb una certa precisió si no es vol obtenir una xarxa on el camí d'anada i de tornada siguin diferents.

4.4. Escenari 3: BGP amb tres ASs

Durant els dos escenaris anterior s'ha comprovat el funcionament del BGP entre ASs veïns però no se n'ha realitzat cap amb més de 2 xarxes diferents. Per tant, l'únic punt d'estudi en aquest escenari és el BGP com a protocol external (eBGP) i sense tenir en compte quines rutes s'exportaran o s'importaran del protocol interior. Mitjançant la configuració d'aquest tercer escenari es vol comprovar quin efecte té en la distribució de xarxes l'atribut de comunitat. Com ja s'ha comentat, modificant aquest valor es pot limitar la extensibilitat d'una xarxa. Dit d'una altra forma, aquest paràmetre permet assignar la privacitat de cada xarxa notificada als veïns per a que actuïn d'una o altra manera.

Així doncs, l'arquitectura que ha permès comprovar el correcte funcionament d'aquesta funcionalitat és senzilla. Partint de l'escenari anterior, s'ha eliminat l'enllaç troncal de la vlan3 i s'ha modificat la configuració de forma pertinent. El següent diagrama mostra la nova topologia:

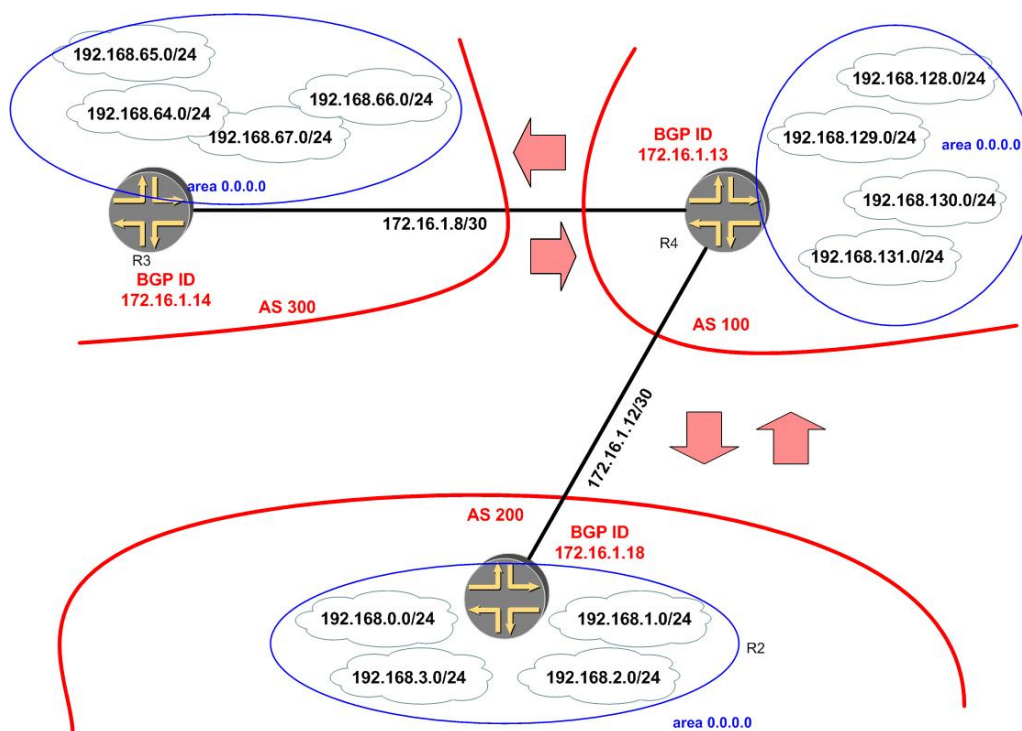


Fig. 4.5 Tercer escenari BGP

La nova arquitectura permet testejar i estudiar els efectes que té un encaminament BGP basat en comunitats. Bàsicament, s'ha anat modificant els valors dels atributs de comunitat per modificar la visibilitat entre diferents sistemes autònoms, o només algunes de les seves xarxes. En el món real, aquests mecanismes són necessaris per poder fer front a restriccions polítiques o simplement administratives. Per al cas de Guifi.net aquests mecanismes

poden aportar més capacitat d'enginyeria de tràfic de la que ja aporten la resta d'opcions del BGP estudiades.

Per mitjà del següent diagrama es pot apreciar la diferencia a l'hora de configurar l'atribut de comunitat amb els valors "no-export" i "no-advertise":

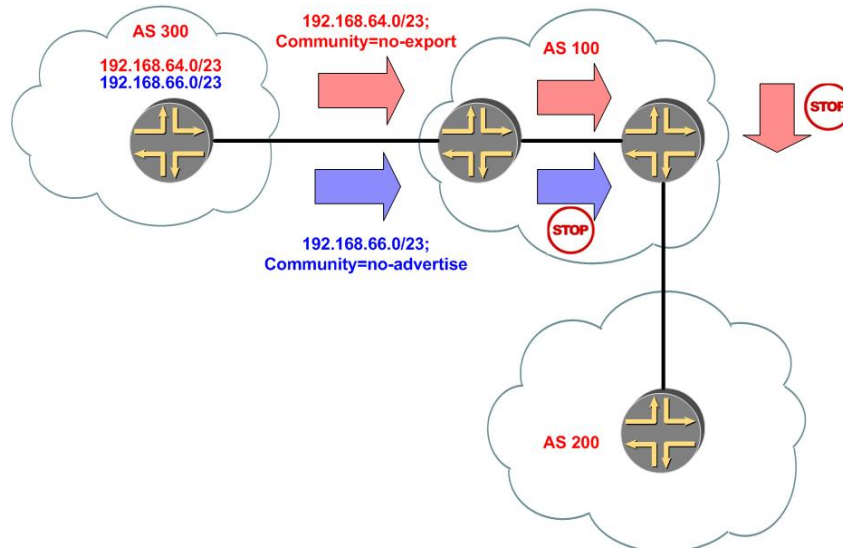


Fig. 4.6 Efecte que té l'atribut de comunitat en la distribució de rutes dins del BGP

Així doncs, i tal i com s'aprecia, els valors restrictius d'aquest atribut permeten limitar fins a quin punt de la xarxa s'anunciarà cada bloc d'adreçament. Es pot veure com l'AS 200 resta aïllat totalment de l'AS 300 independentment del valor. Per altra banda, l'AS 100 si coneix les dues xarxes que el formen, encara que de forma distintiva. Per al cas de la xarxa 192.168.64.0/23 dins de l'AS 100 es veu que serà tractada pel BGP i per tant, anunciada per iBGP als veïns interns. En cap cas serà anunciada pel eBGP als veïns externs. Per altra banda, l'atribut "no-advertise" de la xarxa 192.168.66.0/23 implica que aquesta no serà coneguda pel BGP de l'AS100. Aquesta configuració permet justament que la xarxa no sigui visible o, ho faci a través d'altres protocols. Notar que el primer encaminador de l'AS100 pot afegir la xarxa 192.168.66.0/23 dins de l'IGP i per tant, continuar essent accessible des de la totalitat de l'AS. Per acabar, comentar que el valor "internet" fa que la xarxa al que estigui associat es propagui a tots els veïns i per tant, sigui encaminable de forma global. Durant la realització de l'escenari s'ha vist que aplicar aquest atribut ofereix el mateix comportament que no configurar-ne cap i per tant, es comprova que aquest és el valor per defecte. Tanmateix, i en escenaris grans, el valor de l'atribut Next-Hop pren un valor important. Cada AS que suposi només un transit per a les xarxes anunciades ha de modificar el valor del pròxim salt a l'adreça de la interfície per la qual envia l'anunci fora de l'AS. Fàcilment es pot veure que aquesta acció és per mantenir la connectivitat IP de la xarxa final en l'AS entrant.

La conclusió més important d'aquest escenari és que els atributs s'han de tenir en compte a l'hora de configurar cada punt de la xarxa, donat que aquests afecten directament a la visibilitat i extensibilitat de cada una d'elles.

Arribat aquest punt es veu que el BGP és un protocol que pot aportar beneficis a Guifi.net donat que les funcionalitats son útils i alhora vàlides. L'únic punt que queda per resoldre és quins són els mecanismes d'extensibilitat que ofereix BGP i si són aplicables en aquest cas d'estudi. Guifi.net és una xarxa gran i per això cal veure amb anticipació (abans d'utilitzar el protocol) si el BGP és suficientment extensible.

4.5. Escenari 4: extensibilitat de BGP

Vist que el protocol BGP aporta unes funcionalitats molt concretes i alhora útils per a l'encaminament inter-domini es pot pensar en utilitzar-lo dins de Guifi.net. Abans de fer-ho, però, cal veure quina és la seva extensibilitat i predir si el límit implica alguna restricció important per a Guifi.net. Així doncs, aquest objectiu serà cobert per aquest últim escenari.

Quan s'han nombrat les diferències entre iBGP i eBGP s'ha comentat que en el cas del primer, els prefixos advertits per un peer iBGP mai seran retransmesos a un altre peer iBGP. El motiu és que dins l'iBGP l'atribut "AS-Path" no es modifica i per tant no hi ha mecanisme per detectar un bucle d'encaminament.

La restricció aquí comentada té un efecte directe en l'extensibilitat del BGP, ja que es requereix una malla completa de tots contra tots dins l'iBGP per a la correcta propagació de les rutes i sincronització del protocol. Encara que no tots els nodes d'una xarxa tenen que executar el BGP, la malla requereix una quantitat de connexions entre els encaminadors molt elevada. Concretament, una malla de n nodes requereix un total de $(n \times (n-1))/2$ sessions. Com més augmenta el nombre de sessions més recursos seran necessaris i es generarà més overhead, també en la xarxa.

També es considera important veure que a més quantitat de sessions BGP major quantitat de veïns. Per tant la quantitat de rutes no es veu afectada, però si la quantitat de camins (Path's) per a cada una. Un cop més, els requeriments de recursos són majors.

Per tant, i com es pot veure existeix un problema d'extensibilitat i per tant, cal veure i estudiar les solucions existents. Donat que el nombre d'encaminadors disponibles és baix perd sentit la idea de configurar escenaris pròxims a situacions reals aplicables a Guifi.net, encara que si s'aprofita per testejar les solucions a petita escala i en entorns reduïts.

4.5.1. Reflectors de rutes

El mecanisme de reflectors de rutes (RR) està especificat en l'RFC2796 [7] i defineix el mecanisme pel qual un node BGP entrecanvia informació amb veïns interns (iBGP), a diferència de tota explicació donada fins al moment. Notar, que la única informació generada i intercanviada és el millor camí a cada xarxa i no la taula completa d'encaminament o base de dades de camins.

El funcionament d'un reflector de rutes és senzill d'explicar. Si el millor camí és après d'un client (iBGP) s'informa a la resta de clients i també als no-clients (eBGP). Per altra banda, si el millor camí l'anuncia un no-client (eBGP) es notifica també als clients (iBGP). En aquest últim cas, el camí també es notifica a la resta de nodes eBGP a conseqüència del funcionament habitual del BGP. Tots els encaminadors que mantenen una sessió iBGP amb cada RR s'anomenen clients. En les xarxes grans es recomana utilitzar una configuració amb dos o més nivells de reflectors [8]. Aquesta recomanació parteix de la idea que un reflector d'un nivell és client del reflector del nivell superior, i així successivament. Finalment comentar, que cada grup de clients d'un reflector i el mateix node reben el nom de clúster. En la següent figura es mostra en funcionament dels reflectors de rutes i un exemple de disseny jeràrquic en dos nivells. A més, dins la ramificació de l'esquerra del segon nivell s'ha configurat un disseny amb redundància de RR, comentat tot seguit.

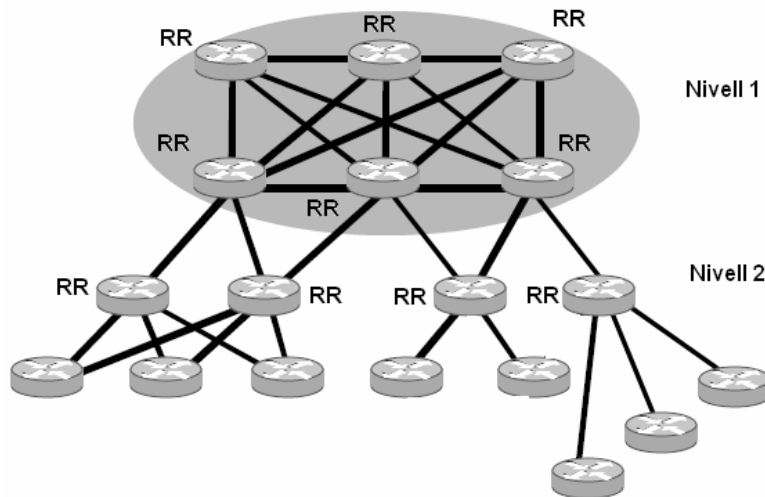


Fig. 4.7 Funcionament dels reflectors de rutes

Normalment, cada clúster o grup de peers tindrà associat un únic reflector. En aquest cas, el clúster estarà associat a l'identificador de l'encaminador amb rol de reflector. Notar que aquesta tècnica genera un punt únic de fallida i per tant s'ha de pensar en algun mecanisme que aporti redundància. Per solucionar aquesta situació i permetre més d'un reflector per cada grup existeix un segon identificador. El segon identificador s'anomena `CLUSTER_ID` i és necessari configurar-lo amb el mateix valor en tots els nodes amb rol de reflector. Durant l'intercanvi d'informació, un reflector descarta tota la informació de qualsevol altre reflector amb el mateix identificador de clúster.

Per a la detecció de bucles d'encaminament s'utilitzen dues variables addicionals, l'identificador de l'origen (`ORIGINATOR_ID`) i la llista del clúster (`CLUSTER_LIST`).

Per a un bon funcionament dels reflectors es recomana seguir una topologia lògica que reflecteixi la física, evitant la creació de cercles tancats. Tanmateix, per al funcionament del BGP és necessària una malla completa de sessions entre tots els reflectors. Notar però, que la quantitat de sessions que s'han de

mantenir alhora és molt menor, permetent una escalabilitat major. De fet, aquest és el mecanisme que dota al BGP d'una major escalabilitat.

4.5.2. Confederacions

El mecanisme de “Confederacions per al BGP” està descrit en l’RFC3065 i especifica una extensió per augmentar l’escalabilitat al BGP i millorar la gestió de la política d’encaminament [9].

La idea principal de l’opció de confederacions es basa en dividir un gran AS amb gran quantitat de nodes BGP en dominis més petits. Aquesta tècnica redueix la quantitat de sessions BGP intra-domini que s’han de mantenir i distribueix la càrrega cap a les sessions inter-domini. En el món real, els ASs resultants de la divisió d’un AS tenen assignats identificadors falsos, ja que la malla de sessions eBGP amb aquests “fake-AS-ID” queda aïllada dins el mateix AS.

Inicialment es pot pensar que al particionar la xarxa en més unitats administratives s’incrementarà la quantitat d’informació d’encaminament, per exemple en els atributs AS_PATH. Aquesta afirmació no és certa, ja que en totes les sessions eBGP externes a l’AS només s’anuncia l’identificador d’AS de la xarxa completa i per tant, la informació d’encaminament es veu reduïda. Com a conseqüència, la gestió de l’encaminament dins l’AS augmenta en complexitat i requereix modificacions cada cop que la topologia canvia. Altrament, qualsevol canvi dins l’AS no es veu reflectit en cap moment en l’eBGP cap a altres ASs i per tant també es minimitza l’overhead en la xarxa.

Per tant, el mecanisme de confederacions al BGP permet unir diversos ASs sota una única nova unitat administrativa o AS que serà la xarxa vista pel BGP de forma global. La següent figura mostra la operació d’aquest mecanisme.

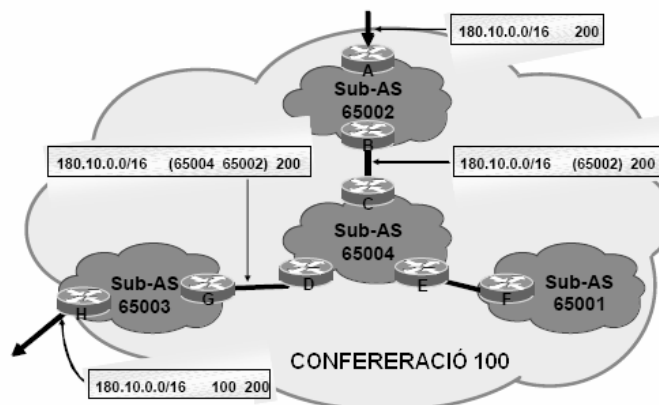


Fig. 4.8 Exemple de confederació formada per 3 ASs

En la figura es veu com l’AS o confederació 100 rep un anunci per a la xarxa 180.10.0.0/16 del seu veí, l’AS (o confederació) 200. Dins la mateixa confederació 100 aquesta xarxa es distribueix per tres sessions eBGP entre els ASs interns. Notar que en l’atribut del camí apareixen els identificadors dels

Sub-AS o ASs interns. Però quan aquesta xarxa 180.10.0.0/16 s'anuncia per eBGP cap als altres ASs només s'adjunta l'identificador de la confederació. Notar doncs, com la quantitat d'informació i complexitat s'ha reduït fora de la xarxa a canvi de més complexitat de gestió en l'interior. Encara que no és l'objectiu principal, també es pot utilitzar el mecanisme de sumari per les xarxes anunciades a l'exterior si és possible, reduint encara més les taules d'encaminament.

Aquest mecanisme es considera que dota al BGP d'una escalabilitat molt elevada, encara que no tant com el cas dels reflectors de rutes. També s'ha de dir, que la complexitat d'implantació és molt major que en el cas dels reflectors. Tanmateix, es considera interessant la seva capacitat per particionar la xarxa en sub-dominis BGP sense necessitar de disposar d'identificadors d'AS d'assignats per les autoritats pertinents. Podria ser interessant el posterior estudi d'aquest mecanisme per tal que els diversos ASs que formen Guifi.net es puguin ajuntar en una única confederació per parlar BGP amb altres xarxes o iniciatives semblants.

4.5.3. Servidors de rutes

La tercera alternativa que es comenta en aquest estudi està descrita en l'RFC1863 [10] i consisteix en l'especificació d'un mecanisme que permet minimitzar la quantitat de sessions BGP i també complexitat de configuració en la majoria dels nodes de la xarxa. Aquest mecanisme s'anomena servidor de rutes i consisteix, a grans trets, en un node que s'encarrega de gran part de la càrrega computacional mantenint una vista de cada node de la xarxa i generant i enviant-li les taules i camins BGP. Així doncs, la quantitat de sessions BGP es veu reduïda a una per a cada sistema autònom cap al servidor de rutes.

Les implementacions actuals d'aquest mecanisme son varies. Aquesta funcionalitat es pot trobar en productes comercials com Cisco fins en implementacions de lliure distribució, com és el cas de Quagga.

Així doncs es pot veure que aquest mecanisme pot ser d'interès en la configuració de punts neutres o d'intercanvi de tràfic entre diferents AS, però no es considera un mecanisme d'utilitat donada la complexitat del seu model [11] i la centralització que es fa de la intel·ligència de la xarxa. Aquesta no es pot aplicar en entorns tant distribuïts com Guifi.net.

4.6. Consideracions de rendiment

Igual que s'ha realitzat en el cas del protocol OSPF també es realitza un estudi més detallat sobre les implicacions del BGP en la càrrega dels equips que l'executen i quins son els requeriments computacionals. Un cop més, s'entén per requeriments computacionals els cicles de rellotge del processador, la memòria i l'ample de banda que consumeix el protocol.

En quant al consum de processador i ample de banda, en el cas del BGP tenen una dependència molt alta. Immediatament després de la configuració BGP tots els veïns intercanvien la informació d'encaminament. Així doncs, la quantitat d'informació d'encaminament que s'intercanvia cada parell de veïns és $O(N + M \times A)$ on N és la quantitat de xarxes, M és la longitud de la xarxa en número de sistemes autònoms i A la totalitat de sistemes que formen la xarxa. Després de l'intercanvi inicial d'informació la quantitat d'overhead enviat a la xarxa depèn només de l'estabilitat d'aquesta. Si no hi ha cap canvi en la xarxa l'únic consum és degut a l'enviament dels missatges de manteniment de la sessió BGP (5 bits/s). La carrega computacional associada a aquest procés és negligible. Si denotem amb C el nombre de canvis per segon que hi ha dins la xarxa la quantitat d'informació addicional que s'ha d'injectar en el pitjor cas serà $O(C \times A)$.

Un cop finalitzat el que es podria anomenar com estat transitori del protocol la utilització d'ample de banda i cicles del processador depenen únicament de l'estabilitat de la xarxa. Tanmateix, quan es diu estabilitat de la xarxa es refereix a totes les xarxes anunciades pel BGP. Per tant, en el cas de la Internet l'estabilitat del BGP depèn de cada una de les xarxes que el formen.

També es considera interessant realitzar alguna diferenciació entre el consum d'ample de banda o temps de processador entre iBGP i eBGP. Mentre en el cas de l'iBGP només es realitza un intercanvi complet de la informació durant l'estat transitori, en el cas del eBGP aquest es realitza de forma periòdica. Normalment, aquest intercanvi es realitza cada 3 minuts. Per tant, es pot assumir que, en el cas de moltes inestabilitats en l'iBGP, els requeriments computacionals que requerirà el protocol seran en el pitjor cas els de l'eBGP durant l'estat transitori o després de la inicialització.

Per quantificar els requeriments de memòria del protocol BGP, en el pitjor cas, és necessari definir les següents variables: es denota N com el nombre de xarxes existents i M com la distància d'ASs coneguda per l'encaminador. Per altra banda, A és la quantitat de xarxes que componen la Internet i K el total de nodes BGP de l'AS en concret. Doncs, els requeriments de memòria es poden aproximar per $O((N + M \times A) \times K)$. Durant aquest estudi s'ha comprovat que els requeriments de memòria si es vol executar el BGP augmenten considerablement respecte en cas de l'OSPF. Durant la realització dels escenaris s'ha comprovat que el procés BGP pot consumir fins a un 20% de la memòria d'un WRT54GL i per tant, i en el millor dels casos resta un 80% per fer créixer la taula d'encaminament. Realitzant comparacions amb alguns resultats experimentals [12] i seguint l'expressió anterior es pot realitzar una aproximació dels requeriments de memòria en situacions ben definides a Guifi.net. Per exemple, si fixem el nombre de xarxes de Guifi.net a 5000 (usuari més troncs), amb una distància coneguda de 10 ASs, sense realitzar cap tipus de peering amb el BGP global de la Internet ($A=N=2000$) i un total de 50 nodes executant el BGP s'estima una taula d'encaminament de fins a 2Mbytes. Tot i que aquesta xifra pot ésser una estimació no resulta un consum del 80% de memòria però sí és elevat. Per tant, cal pensar en utilitzar els mecanismes per a filtrar rutes (prefix-list, analitzats durant aquest estudi) entrants o sortints a la base de dades del BGP. Aquest mecanisme, entre d'altres, permetrà acotar

la mida de la taula d'encaminament sense que suposi un empitjorament en el funcionament de la xarxa. Durant aquest estudi s'ha comprovat que els requeriments de memòria si es vol executar el BGP augmenten considerablement respecte en cas de l'OSPF i per tant és interessant minimitzar el nombre de nodes que executin ambdós protocols.

Adicionalment es comenten un parell d'aspectes més del BGP focalitzats a controlar o reduir consum de recursos del protocol. El primer és el definit en l'RFC2439 i implementat en el cas de Quagga i Bird. A grans trets es pot comentar que es defineixen dos llindars, a través dels quals l'encaminador accepta el creixement de la taula d'encaminament (llindar inferior) o comença a descartar rutes (llindar superior). La definició inicial d'aquest mecanisme és per acotar el desastrós efecte que té en la Internet el comportament inestable d'alguna xarxa amb qualsevol tipus de problema. Aquest mecanisme és utilitzat en les grans troncs de la Internet i no és considera necessari en el cas de Guifi.net. Tanmateix està implementat i cal considerar-lo si Guifi.net executa eBGP cap a la Internet. El segon, i possiblement més útil a curt o mitjà termini és l'existència del comptador `MinRouteAdvertisementInterval`, pel qual es defineix el temps mínim pel qual s'informarà d'una o totes les xarxes a un peer en particular. Cal estudiar amb detall la modificació d'aquest paràmetre en entorns reals, ja que té un efecte directe en la convergència del BGP i per tant, la durada de l'estat transitori, durant el qual es consumeixen més recursos computacionals i de xarxa.

4.7. Conclusions

Abans de donar una solució final o proposta d'implantació a Guifi.net s'ha mantingut un punt dedicat a l'extracció de conclusions fruit de l'estudi del protocol pròpiament dit.

Durant aquest estudi, s'ha pogut comprovar que el BGP és un protocol molt complex i amb uns requeriments computacionals més elevats que l'OSPF. Tanmateix, el disseny de la xarxa i de les polítiques són elements que s'han d'estudiar amb gran detall si es vol aconseguir una xarxa estable i controlada. A més, el protocol és altament escalable si es compten amb alguns dels mecanismes com són els reflectors de rutes o les confederacions.

Fruit de la part pràctica cal dir que la configuració és molt complexa donada la necessitat d'utilitzar llistes de control d'accés i mapes de ruta (route-map), a més de realitzar les accions pertinents amb els atributs de cada ruta per a què sigui propagada de la forma que es desitgi. Per a una simplificació del tractament de la configuració es recomana altament l'ús de Peer-Group's. Aquesta condició agrupa els veïns per aplicar-los una mateixa política de BGP, sense necessitat de realitzar una configuració tant extensa com peers BGP.

Per últim, incidir un cop més que el BGP és un protocol mol complex d'administrar així que cal anar amb especial atenció amb la creació de bucles o forats negres d'encaminament, ja que són difícils d'acotar i solucionar.

CAPÍTOL 5. APLICACIÓ A GUIFI.NET

5.1. Introducció

Aquest punt pretén ésser una extensió de les conclusions associades a cada protocol, però amb un punt de vista diferent. L'objectiu principal d'aquest capítol és la redacció d'una proposta de disseny per a Guifi.net. Aquest nou disseny es basa en visió de Guifi.net com una rèplica de l'actual Internet i per tant, amb necessitats d'encaminament i assignació d'IP, entre d'altres, ben diferenciades. Per tant és necessari tractar la gestió de Guifi.net des del punt de vista de múltiples sistemes autònoms que executen un protocol d'IGP intern i un EGP per comunicar-se amb la resta de sistemes. Els protocols inicialment utilitzats són els estudiats en aquest treball, OSPF com a IGP i el BGP com a EGP.

S'ha vist que disseny actual de Guifi.net no és sostenible i la proposta que es realitza aquí ha de ser robusta i escalable al creixement notori de Guifi.net. A l'hora de realitzar aquesta proposta s'ha decidit dividir aquest punt en dos apartats ben diferenciades. La primera part correspon al disseny de la xarxa troncal, que executarà el protocol BGP en una sèrie de nodes per proveir connectivitat inter-domini i en alguns casos també intra-domini. La segona part descriu la configuració proposada de la xarxa d'accés o distribució, donat que Guifi.net és una xarxa inalàmbrica que dona accés als seus clients. Aquesta part executarà el protocol OSPF i evidentment inclourà la gran majoria de l'equipament que conforma la infraestructura de Guifi.net. El protocol OSPF com a protocol d'estat d'enllaç és una bona elecció donades les característiques canviants del canal ràdio.

Finalment és troba una nota aplicable a la implantació d'aquesta o qualsevol altra reforma a Guifi.net. Tot procés de migració és altament costós i delicat i per tant, ha d'ésser degudament planificat i executat.

5.2. Disseny de la xarxa troncal

Aquest document proposa la idea de veure Guifi.net com una rèplica en miniatura de la Internet i per tant aplicar les mateixes polítiques de disseny. Tanmateix, donada la distribució de Guifi.net i el seu caràcter de xarxa oberta sense fils dificulta l'existència de creació de punts neutres d'intercanvi de tràfic. Aquests descarreguen la utilització dels enllaços troncal eliminant part del tràfic provinent del trànsit entre els diferents AS. Aquesta diferència no impedeix que el protocol BGP sigui el candidat idoni per donar solució a l'encaminament inter-domini, encara que és previsible la necessitat de xarxes troncal d'alta capacitat per compensar l'efecte que s'ha comentat. Així doncs, i com a primera aproximació, cada comarca es pot plantejar com un sistema autònom. En el cas de comarques amb gran densitat de nodes podria pensar-se en tractar-ho com 2 ASs diferents, encara que una altra alternativa pot ésser alguns dels mecanismes d'escalabilitat del BGP.

5.2.1. Disseny global de Guifi.net

El disseny de la xarxa BGP de Guifi.net es pot basar per dos elements diferents i que alhora que en conformen la xarxa també en defineixen la seva topologia i limitacions. El primer d'aquests elements és la disponibilitat d'equips amb capacitats hardware més elevades que els WRT54GL, anomenats "supertrastos" o, donada l'escassetat d'aquests, també es pot considerar els nodes que puguin assegurar una bona disponibilitat i connectivitat. El segon és la pròpia configuració i disponibilitat dels enllaços de llarga distància, normalment coberts amb canals a 5GHz o altres tecnologies emergents.

La següent figura mostra a mode d'exemple com es pot dividir Guifi.net en diferents sistemes autònoms prenent com a líndars les divisions comarcals del territori. Aquesta és una possible política de disseny, encara que no la única.

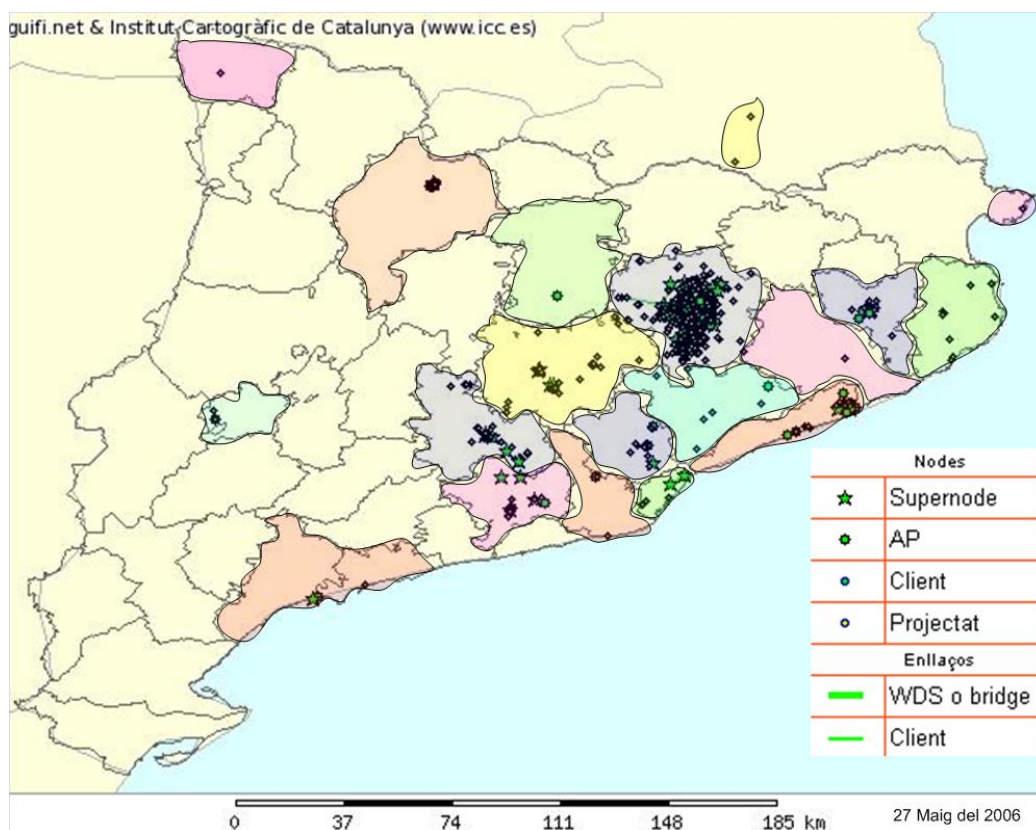


Fig. 5.1 Configuració de Guifi.net amb diferents ASs

En la figura s'hi mostren tots els nodes que formen Guifi.net. Cal distingir però, les diferències entre un supernode o un AS i els clients, que mostraran un comportament més inestable. Tanmateix, en la comarca de l'Osona hi ha gran part dels nodes i pot ser interessant tractar-ho com diversos ASs diferents.

Per donar a la xarxa de la suficient capacitat per tractar tot el tràfic de forma eficient pot ser interessant realitzar un estudi avançat de com es redistribueixen les rutes i en conseqüència, actuar sobre els atributs de cada xarxa per aconseguir el tracte i encaminament desitjat. Entre ASs adjacents es pot estudiar la possibilitat d'anunciar els rangs sense sumaritzar o amb una

sumarització pobra, assignant diferents valors de l'atribut MED en funció de cada punt de sortida. Així es pot distribuir eficientment la càrrega d'entrada a un AS des de l'adjacent si existeix més d'un enllaç entre ells. Tanmateix, també es recomana anunciar el rang sumaritzat de tot l'AS, amb l'atribut de comunitat amb valor públic, i que serà propagat per tota la xarxa. Tal i com s'ha comentat en l'escenari tercer, el valor de l'atribut Next-Hop s'ha de modificar cada cop que l'anunci surti d'un AS. Notar doncs, que per un AS no veí és independent el punt d'entrada a l'AS destí i per tant es pot aprofitar la sumarització. Per al cas dels rangs no sumartizats es recomana anunciar-los amb l'atribut de comunitat amb un valor no-export per tal de limitar el seu abast dins l'AS veí.

La configuració d'enllaços redundants també pot ésser interessant si es volen minimitzar els talls de connectivitat encara que s'entén que la seva disponibilitat és molt baixa o inexistent.

Per altra banda i fruit de l'estudi dels mecanismes d'escalabilitat al BGP també es pot donar una solució a la possible interacció de la xarxa Guifi.net amb la d'altres iniciatives semblants, institucions o qualsevol xarxa fora dels seus límits administratius. La millor solució arribada la situació que Guifi.net interactua amb altres xarxes és la de configurar tots i cadascun dels ASs que la conformen com una única confederació BGP. Donada aquesta situació, Guifi.net independentitza tota la seva topologia interna i assignació d'adreces de les xarxes a les quals connecta, anunciant únicament una única unitat administrativa amb un rang d'adreces, corresponent a la sumarització de totes les xarxes. La figura següent mostra aquesta situació aplicada al cas on Guifi.net es connecti per exemple amb les xarxes de la Universitat Politècnica de Catalunya (UPC, AS100) i la Fundació i2CAT, AS 200. Tanmateix, també es pot executar el BGP amb altres iniciatives Wi-Fi a Catalunya o fora del principat.

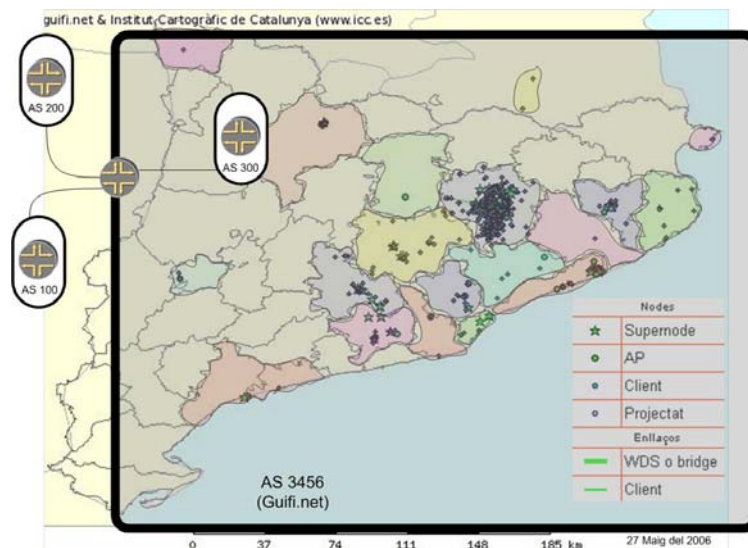


Fig. 5.2 Configuració de Guifi.net com una confederació

Notar també, que el disseny de BGP no requereix una assignació jeràrquica i per tant aquesta proposta no implica un canvi complet en l'assignació d'aquestes.

5.2.2. Disseny de cada sistema autònom

És evident que en la majoria dels sistemes autònoms existiran més d'un node que executi el BGP i per tant també és important realitzar un bon disseny del BGP com a protocol IGP de Guifi.net. Encara que aquest no és l'encarregat de l'encaminament intra-domini, si s'han d'injectar les rutes (o algunes, o cap) dins del mapa de la xarxa interna. Tanmateix, ja s'ha comentat que tots els nodes iBGP d'un AS han d'estar enllaçats formant una malla completa o utilitzar alguna altra solució. La figura que es mostra a continuació planteja una primera aproximació al disseny de l'iBGP per tal de mantenir el protocol BGP funcional, però no com a protocol encarregat de l'IGP dins de l'AS.

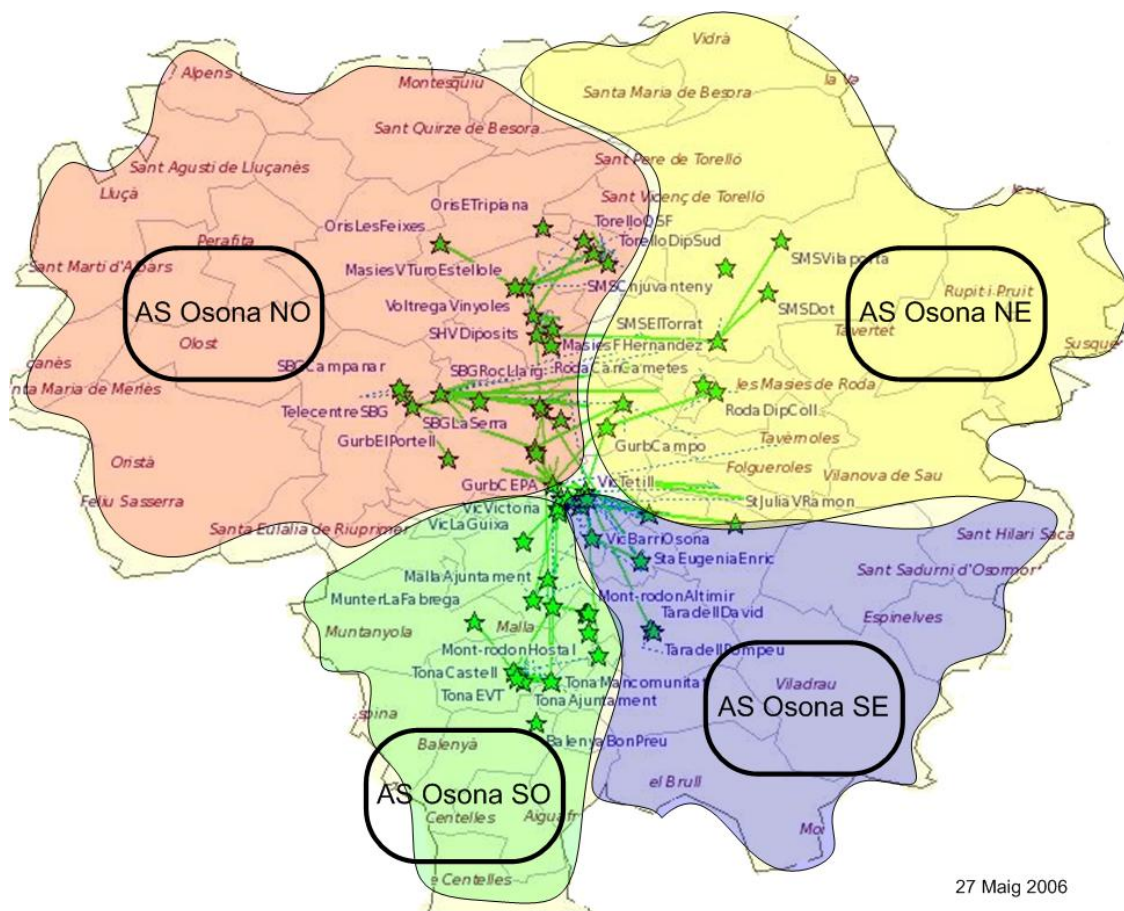


Fig. 5.3 Configuració d'un AS amb gran quantitat d'enllaços frontera

S'ha aprofitat aquesta figura per presentar la situació comentada anteriorment, on l'Osona es pot tractar com 4 ASs diferents. Notar que en la figura s'han exclòs tots els punts d'accés i clients, ja que aquests són els que executaran el protocol OSPF.

Per tal de distribuir la càrrega de les troncales s'han aprofitat tots els enllaços entre diferents AS i albergar-hi tràfic inter-domini. Per tant, apareixeran gran quantitat de nodes BGP dins del sistema autònom que seran punts d'entrada i sortida i hauran d'executar l'eBGP i l'iBGP. En conseqüència, la malla de sessions BGP pot resultar una càrrega computacional considerable i que afecti al rendiment del propi equipament. Plantejada aquesta situació és interessant

considerar l'utilització d'algun mecanisme d'escalabilitat per tal de minimitzar o distribuir la càrrega computacional.

Com a solució al problema d'escalabilitat és possible pensar en la idea de configurar un o diversos reflectors de rutes BGP, encarregats de mantenir la malla iBGP reduïda i alhora consistent. En els casos més extrems o de sistemes autònoms molt grans existeix la possibilitat de realitzar un disseny en dos nivells, tal i com s'ha explicat en el quart escenari del protocol. Per altra banda, es podria plantejar l'opció d'utilitzar un servidor de rutes, també disponible en Quagga/Bird. Aquesta altra opció no es presenta tant atractiva ja que es centralitza la dependència del funcionament de l'iBGP en un únic node. A més, els requeriments computacionals d'aquest servidor son més elevats dels que pot oferir un WRT54GL o semblant.

Arribat aquest punt és interessant realitzar una estimació de quantes i quines rutes formaran la taula BGP de cada node o AS. Separem aquestes rutes en les rutes generades pel propi AS ($MyBGP_RT$) i les rutes que han estat injectades per cada AS (BGP_RT) a la taula global del BGP.

Les rutes generades pel propi AS, o $MyBGP_RT$ corresponen inicialment, a la quantitat de rutes generades per la màxima sumarització de l'adreçament assignat a l'AS. Notar que idealment aquesta ruta ha de ser 1. Aquesta ruta és la que s'importa al BGP amb un valor de comunitat amb valor "públic" per a que sigui retransmesa per tota la xarxa de forma global. Denotem aquesta quantitat de rutes com $f(IPs)$. Addicionalment a les rutes sumaritzades de tot l'AS s'han d'afegir les rutes de l'iBGP en cas que l'AS estigui format per més d'un node BGP. Notar que la quantitat de rutes oscil·la entre la quantitat d'enllaços BGP (millor cas) o la quantitat d'enllaços o xarxes OSPF. Notar que depèn exclusivament de la topologia interna i el nivell de sumarització utilitzat. Denotem aquest valor com $f(iBGP)$.

Les rutes dels altres ASs, o BGP_RT també es poden dividir en dos si es pensa que l'impacte en la taula d'encaminament és molt diferent. Les primeres corresponen a la dels ASs remots i que per tant, no existeix una sessió BGP. Notar que la quantitat de rutes correspon al resultat de $f(IPs)$ aplicat a cada AS i per tant, es pot especificar com

$$\sum^M f(IPs_i) \quad (5.1)$$

on M correspon a la quantitat d'ASs de tot Guifi.net. Les segones rutes correspondran a les generades pels ASs colindants i que estaran etiquetades amb l'atribut de comunitat limitant la seva extensibilitat. En aquest cas, si existeix sessió BGP entre els ASs o peers. Si la només existeix un enllaç entre els ASs la ruta ja esta considerada en la part anterior. Però, si existeix més d'un enllaç entre els ASs, i havent recomanat realitzar balanceig de càrrega a nivell BGP cal fer una última consideració. Per al balanceig cal advertir les xarxes no sumartitzades per tal de diferenciar els enllaços. Notar que en el pitjor cas, on no es realitza cap tipus d'agrupació o filtrat de les xarxes que s'anuncien aquest

valor correspon al resultat d'aplicar $f(iBGP)$ a cada AS colindant. Per altra banda, en el cas òptim la quantitat de rutes es podrà sumaritzar fins a un valor equivalent al d'enllaços entre ambdós ASs. Per tant, si prenem el pitjor cas, la quantitat de rutes dels ASs colindants es pot especificar com

$$\sum^{M'} f(iBGP_j) \quad (5.2)$$

on M' correspon a la quantitat d'AS colindants. Per tant, prenent les dues expressions s'obté que la mida de la taula d'encaminament BGP d'un node qualsevol es pot aproximar per

$$BGP = BGP_RT + M_y BGP_RT = \left(\sum^M f(IPs_i) + \sum^{M'} f(iBGP_j) \right) + f(iBGP) \quad (5.3)$$

si es prenen les pitjors condicions de disseny especificades en cada secció. En la part independent al propi AS (BGP_RT) es pot apreciar que el primer terme pren un pes més important si Guifi.net es subdivideix en gran quantitat d'ASs. El segon terme pren importància a major quantitat d'enllaços múltiples entre ASs veïns i es veu directament afectada per l'efectivitat de les polítiques de d'anunci de xarxes internes al BGP. L'últim terme reflecteix l'impacte de la pròpia topologia de l'AS.

5.3. Disseny de la xarxa de distribució

En aquest punt és realitzen un seguit de recomanacions a l'hora de la configuració de la xarxa interna de cada AS, fruit de l'experiència d'aquest estudi en el disseny i planificació de xarxes i la corresponent aplicació a Guifi.net. Donat que la xarxa està dividida en dominis, cada AS requereix una configuració de l'OSPF independent de les altres. Per tant, i com ja s'ha comentat, l'assignació de direccions a les interfícies dels enllaços troncal de cada xarxa és independent de la resta.

Donat que l'OSPF és un protocol d'enllaç eficient en captar les característiques dels enllaços i adaptar les taules d'encaminament es considera interessant la idea de realitzar un disseny de les àrees dependent de les característiques locals de cada zona. Aquest comentari vol dir que si per exemple existeix una zona ampla i plana, possiblement no hi haurà problemes amb els canals ràdio propers. Aquesta situació afavoreix la creació d'una àrea gran, ja que tots els canals presentaran una bona disponibilitat i es produiran pocs canvis en la topologia. Com més gran sigui l'àrea major és, previsiblement, la sumarització que és pot configurar en els seus nodes ABR. Per altra banda, una zona on existeixin gran concentració d'altres xarxes, o per exemple un enllaç problemàtic és recomanable configurar una zona aïllada. Aquesta configuració minimitzarà les pèrdues de connectivitat de les xarxes adjacents i la quantitat d'informació que s'enviarà per tota la xarxa.

La següent figura mostra un exemple de proposta de configuració de la xarxa OSPF d'un AS, concretament l'AS Osona NO. Notar que s'ha configurat una àrea troncal gran. Totes les àrees que interconnecten amb altres ASs (a excepció de la backbone) obtenen el rol de NSSA i la resta compten amb més d'un encaminador ABR. Partint de la situació en que cada AS rep un únic rang d'adreçament i en el cas que siguin diversos siguin consecutius la política d'assignació d'adreces dins cada AS ha de ser el més jeràrquica possible.

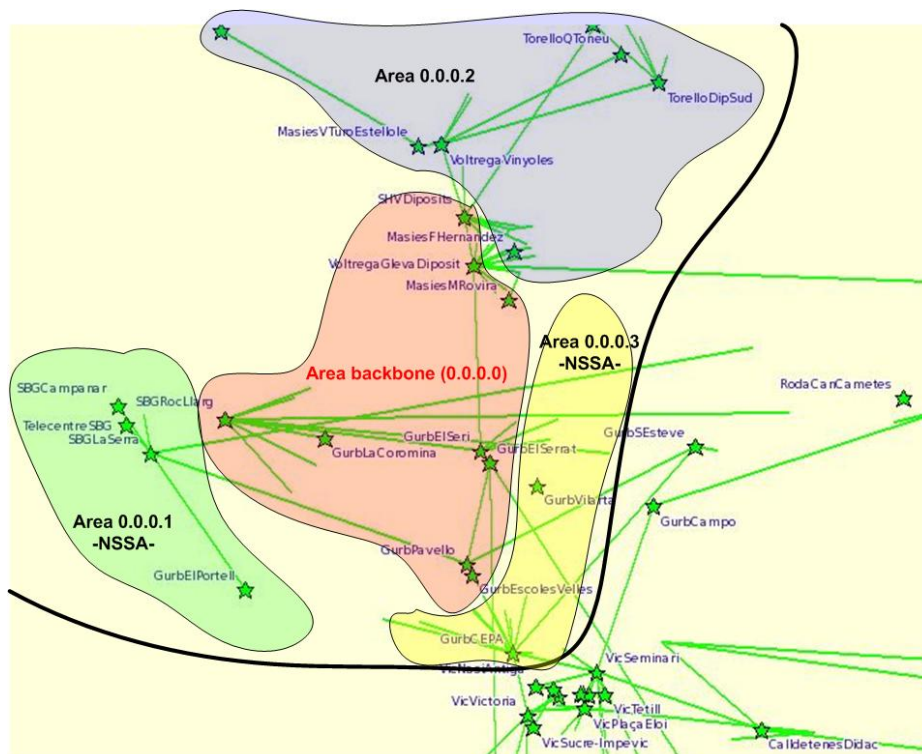


Fig. 5.4 Xarxa interna de l'AS

Deixant de banda les consideracions específiques de cada zona i que, afectaran en gran mesura al seu rendiment, es poden prendre una sèrie de mesures com a punts de disseny aplicables a cada una de les xarxes OSPF:

- Un backbone gran i pròxim als enllaços externs de l'AS.
- Si l'AS és molt gran, una sèrie d'àrees al voltant del backbone. És important que hi hagin diversos encaminadors ABR dins de cada i cap a cada àrea. Aquestes àrees poden constituir la major part de la xarxa de distribució.
- Prioritzar la configuració d'àrees stub si es realitza des d'un node amb bona disponibilitat o on només es disposi d'un únic enllaç per cobrir una zona. Recordar que tot l'encaminament fora de l'àrea es pot basar en una única ruta per defecte i per tant, aquesta mesura resulta un impacte molt gran a l'hora de minimitzar la taula d'encaminament.

- Es pot pensar en separar els nodes que executin BGP en àrees NSSA. Aquesta mesura pot ser interessant a l'hora de separar el troncal BGP (on hi haurà tràfic inter-domini) del troncal OSPF (encarregat només de tràfic inter-domini).
- Maximitzar l'ús dels sumaris de l'OSPF per minimitzar les taules d'encaminament. Per aconseguir aquest punt és molt important una assignació d'adreces el més jeràrquica possible.

Un punt que cal tenir en compte és com es distribuiran les rutes del BGP dins l'OSPF. Si l'AS està format per un únic node BGP no és necessari injectar les rutes del BGP dins del procés OSPF si es genera una ruta per defecte i aquesta es distribueix cap a l'OSPF. En canvi, si l'AS està format per més d'un node i per tant, existeix redundància inter-domini si és necessari injectar aquestes rutes dins l'OSPF. En aquesta situació és especialment important configurar cada àrea amb el rol correcte i realitzar una configuració dels sumaris de l'OSPF eficient.

5.4. Implicacions d'implantació

Tot procés d'implantació o migració entranya els suficients dubtes com per arribar a dubtar sobre ella mateixa. Evidentment, és un procés costós i no existeix cap solució universal ni cap manual sobre un procés de migració. En qualsevol cas, es recomana una planificació llarga i específica per a cada zona o AS i una altra de global abans d'iniciar qualsevol modificació. Arribat aquest extrem, i amb fins com els d'obtenir l'arquitectura òptima es pot plantejar l'opció de la simulació abans de la implantació.

Per al cas de Guifi.net la reconfiguració total o parcial de la xarxa IP pot esdevenir el major entrebanc a l'hora de plantejar un disseny com el proposat en aquest document. Tanmateix, augmenta la viabilitat el fet de disposar d'un espai d'adreçament ampli lliure, sobre el qual començar a assignar adreces a la nova xarxa no impliqui la modificació de l'actual.

Partint de la situació en que tot Guifi.net forma part de l'àrea troncal d'una única xarxa OSPF es podria pensar en aprofitar aquesta situació. És possible configurar tots els nodes BGP dins del mateix AS sense que el funcionament de l'OSPF, i per tant Guifi.net, es vegi alterat. En aquesta situació, tot Guifi.net es comportarà com un únic AS. Arribat aquest punt només és necessari anar segregant cada AS final de la xarxa completa. Conseqüentment, l'àrea troncal de cada nou AS ja no formarà part de la part resilient i per tant, ja es podrà configurar com a xarxa interna del nou AS. El mecanisme d'enllaços virtuals per mantenir la connectivitat entre les diferents àrees backbone generades amb la subdivisió de la xarxa. La configuració de cada xarxa OSPF també s'ha de dur a terme amb un grau de planificació elevat.

CAPÍTOL 6. CONCLUSIONS FINALS

Durant aquest projecte s'ha estudiat la situació i entorn de Guifi.net. S'ha vist que aquesta iniciativa té uns punts forts que la converteixen en única en el seu àmbit. Les característiques diferencials són el programari que es desenvolupa per a l'autoconfiguració del nodes i el model que segueixen, totalment obert i distribuït. Tanmateix, les associacions lliures d'usuaris son tantes i tant nombroses que comencen a esdevenir una alternativa real a les opcions comercials avui existents.

Mitjançant les experiències pràctiques s'ha pogut comprovar que els protocols d'encaminament són molt robustos i adaptats als objectius pels quals han estat dissenyats. Una novetat respecte la situació actual de Guifi.net és que en cada nivell de la xarxa s'han d'executar uns protocols ben diferenciats, ja que també son les funcions que es realitzen en cada punt. Per mitjà de la configuració robusta d'aquests protocols es possible crear una xarxa estable i optimitzada. Tanmateix, l'escalabilitat de la xarxa si es segueix un disseny jeràrquic o anàrquic és notablement major. En al cas de Guifi.net aquest factor és crític, davant el fort creixement d'aquesta xarxa i extensió a nous emplaçaments.

Per altra banda, després de realitzar una sèrie d'escenaris s'han obtingut configuracions robustes i una llista dels punts forts a tenir en compte durant el disseny i implantació d'aquestes tecnologies. Un punt a destacar és que s'ha utilitzat el mateix maquinari que Guifi.net i per tant, les maquetes configurades es poden considerar talment com zones concretes de Guifi.net. Tanmateix, i com a última tasca s'han redactat una sèrie de propostes concretes per a Guifi.net per tal de realitzar una configuració vàlida i altament escalable. Notar que al llarg d'aquest estudi s'han tractat totes les opcions de cadascun dels protocols, oferint una visió global i també concreta de cadascun dels protocols. Es creu però, que la llista de propostes és força genèrica a qualsevol zona de Guifi.net i per tant, és possible que algunes de les zones de la xarxa necessitaran un tracte diferencial i configuració més específica.

Durant la realització d'aquest estudi només s'han realitzat modificacions software en un equipament determinat i per tant es creu que el seu impacte ambiental és quasi inexistent. A més, la seva aplicació a Guifi.net no anima activament la implantació de nous enllaços. Aquestes situacions requereixen un estudi ambiental més profund del que es realitza en el present projecte. Tanmateix, sí es vol recomanar una assignació de canals justa i coherent. En les situacions que hagin de coexistir diverses xarxes; que en el cas de tecnologies 802.11 és sovint, cal tenir present que les bandes lliures són d'ús obert a tota la població. Aquesta situació pot ésser tractada entre les bandes implicades per tal d'arribar a una situació més justa, sostenible i, possiblement, eficient.

Abans de finalitzar aquestes conclusions es vol oferir una llista de tasques que es poden iniciar dins de Guifi.net i per tant, conformen les línies futures d'aquest estudi.

En la línia de l'encaminament, i pensant en optimitzar cada una de les configuracions abans d'implantar-les es creu molt interessant la simulació d'aquest protocols en les mateixes condicions, quantitat i topologia de les que s'executaran en l'entorn real. Per a tal objectiu, existeixen simuladors d'ambdós protocols utilitzats, com són Totem [13] i C-BGP [14] i que per tant, ofereixen una plataforma per a fer-ho, podent concentrar la major dedicació a la simulació en comptes de la creació o aprenentatge de l'eina.

Una altra línia ben diferenciada que es pot obrir és la implantació d'Ipv6 a Guifi.net. Actualment, Guifi.net es basa en un encaminament Ipv4 privat, amb les deficiències que en algunes aplicacions pot ocasionar. La utilització del protocol Ipv6 permetria a Guifi.net la configuració d'una xarxa amb adreçament públic i per tant, amb més visibilitat cap a l'exterior. Encara que la implantació d'Ipv6 a nivell global és un incògnita que fa temps va aparèixer i encara no s'ha resolt, si s'ha vist com algunes xarxes ja han iniciat el procés de migració cap al nou Internet.

Un cop la xarxa té un encaminament eficient i més o menys optimitzat, si es vol predir quin serà el comportament de la xarxa davant una situació o una aplicació concreta es pot pensar en dotar-la de mecanismes de qualitat de servei (QoS). La implementació actual de l'OpenWRT ja és compatible amb aquests mecanismes de classificació i tractament del tràfic. Per tant, aquesta proposta esdevé una altra interessant opció, independent de les altres, però alhora complementaria.

L'avaluació del desenvolupament d'aquest estudi és positiva. S'ha seguit la planificació inicial amb petites derives però un resultat final a la baixa. Per ampliar aquesta informació cal referir-se a l'annex 4, Planificació i pressupost del projecte.

BIBLIOGRAFIA

- [1] B. Fortz. *On the evaluation of the reliability of OSPF routing in IP networks*. Institut d'Administration et de Gestion. On-line a URL: <<http://www.poms.ucl.ac.be>>
- [2] V. Eramo, M. Listanti, G. Gasparro, A. Cianfrani. *Optimization in the shortest path First Computation for the quagga software routing*. University of Roma "La Sapienza", INFOCOM Dept. On-line a: URL: <<http://net.infocom.uniroma1.it>>
- [3] A. Retana. *IP Routing Protocol Scalability Theory and Examples*. Cisco Systems, 2001.
- [4] Tatiana B. Pereira and Lee L. Ling. *Network Performance Analysis of an Adaptive OSPF Routing Strategy – Effective Bandwidth Estimation*. International Telecommunication Symposium – ITS 2002, Natal, Brazil
- [5] T. Davis. *Linux Ethernet Bonding Driver mini-howto*. On-line a URL: <<http://www.kernel.org>>.
- [6] A. Ramanath. *A Study of the interaction of BGP/OSPF in Zebra/ZebOS/Quagga*. Department of Computer Science, State University of New York at Stony Brook. On-line a URL: <<http://www.cs.stonybrook.edu>>.
- [7] T. Bates, R. Chandra, E. Chen. *BGP Route Reflection - An Alternative to Full Mesh IBGP*. April 2002. On-line a URL: <<http://www.ietf.org>>
- [8] N. Taft. *The Basics of BGP Routing and its Performance in Today's Interne*, Advanced Technology Labs. May 2001.
- [9] P. Traina, D. McPherson, J. Scudder. *Autonomous System Confederations for BGP*. February 2001. On-line a URL: <<http://www.ietf.org>>.
- [10] D. Haskin. *A BGP/IDRP Route Server alternative to a full mesh routing*. October 1995. On-line a URL: <http://www.ietf.org>
- [11] *Quagga Routing Software Suite Documentation*. Section 10.1. On-line a URL: <http://www.quagga.net/docs>.
- [12] Y. Rekhter. *BGP Protocol Analysis*. October 1991. On-line a URL: <<http://www.ietf.org>>.
- [13] J. Lepropre, S. Balon, G.Leduc. *TOTEM: A Toolbox for Traffic Engineering Methods*. Research Unit in Networking - University of Liège, Belgium. On-line a URL: <<http://www.ieee-infocom.org>>
- [14] B. Quoitin, S. Uhlig. *Modeling the Routing of an Autonomous System with C-BGP*. University of Louvain-la-Neuve. On-line a URL: <<http://www.info.ucl.ac.be>>.

ACRÒNIMS

ABR	Area Border Router
AS	Autonomous System
ASBR	Autonomous System Boundary Router
BGP	Border Gateway Protocol
CPU	Central Processing Unit
eBGP	External BGP
EGP	External Gateway Protocol
EPSC	Escola Politècnica Superior de Castelldefels
FTP	File Transfer Protocol
GPL	General Public Licence
HDLC	High Data Link Control
IAB	Internet Architecture Board
iBGP	Internal BGP
ICMP	Internet Control Message Protocol
IEEE	Institute of Electrical and Electronic Engineers
IETF	Internet Engineering Task Force
IGP	Internal Gateway Protocol
IMAP	Internet Message Acces Protocol
IP	Internet Protocol
IPv6	Internet Protocol versió 6
IRC	Internet Realy Chat
LAN	Local Area Network
LSA	Link State Advertisment
MED	Multi-Exit Discriminator
NAT	Network Address Translation
NSSA	Not So Stubby Area
NVRAM	Non Volatile RAM
OSPF	Open Short Path First
POP	Post Office Protocol
PPP	Point to Point Protocol
QoS	Quality of Service
RAM	Random Access Memory
RFC	Request For Comments
RIP	Routing Internet Protocol
RR	Route Reflector
SMS	Short Message System
SPT	Source Path Tree
SSL	Secure Socket Layer
TCP	Transport Control Protocol
UDP	User Datagram Protocol
VLAN	Virtual Local Area Network
VLSM	Variable Lenght Subneting Mask
VoIP	Voice over IP
WAN	Wide Area Network
WLAN	Wireless Local Area Network

ANNEX 1. PREPARACIÓ DELS WRT54GL

A1.1. Introducció

L'objectiu principal d'aquest punt és mostrar els passos seguits per tenir els WRT54GL totalment funcionals per als escenaris a partir dels models adquirits al proveïdor. De la mateixa forma, tots els passos intermedis, o altres procediments que es consideren una transferència de coneixement també es troben recollits en aquest punt. Per tant, els escenaris presentats al llarg del document han d'ésser fàcilment reproduïbles seguint aquestes pàgines.

A1.1. Elecció de la distribució i versió

Per veure la possible compatibilitat dels aparells adquirits amb les distribucions escollides és necessari veure concretament quin hardware (chipset, firmware, memòria, etc) concret incorporen els aparells. Així doncs, es disposa d'equips amb 4Mbytes de memòria FLASH i 16Mbytes de memòria RAM. A més, el chipset és el model Broadcom 5352 i per tant:

- Els WRT54GL con compatibles amb OpenWRT i
- també són compatibles amb DD-WRT.

Les dos opcions comentades es distribueixen ASIS ("tal i com son") i sense cap suport especial apart de la documentació disponible i la experiència d'altres usuaris.

A1.1.1. OpenWRT

OpenWRT es distribueix en dos formats diferents, en funció del tipus de sistema de fitxers. La primera opció, SpashFS, proporciona un sistema de fitxers mínim i de només lectura. Una segona partició en format JFFS2 s'utilitza per guardar les modificacions realitzades, oferint un sistema també d'escriptura. El principal avantatge és que s'aconsegueix major nivell de compressió. La segona opció consisteix en utilitzar una única partició en format JFFS2.

Per aquest estudi s'escull l'opció SquashFS ja que ofereix la capacitat de restablir el sistema formatejant la partició d'escriptura i alhora, ens ofereix més capacitat per a la instal·lació de programari addicional. A més, cal dir que la versió actual d'OpenWRT, anomenada White Russian RC5, no suporta el mode de recuperació (failsafe mode).

Segons la taula [de hardware](#), el WRT54GL és un WRT54Gv4. La instal·lació de OpenWRT es pot realitzar des de la interfície web d'administració. Aquest

procediment només és útil el primer cop, i amb el firmware que proporciona el fabricant en estat funcional. Per poder flashear els WRT54GL les pròximes vegades és necessari seguir el següent procediment:

- Es connecta un PC Linux al port 1 del WRT54GL. A la interfície WAN del Linksys hi ha d'haver configurada una IP, encara que no cal connectar el port.
- S'inicia un servidor TFTP i s'executa la pujada (PUT) de la imatge (normalment fitxer.bin) a la IP 192.168.1.1. Durant la repetició d'aquest procés s'observa que el procediment es pot realitzar des d'un terminal Windows XP com Linux.
- Es connecta l'alimentació del WRT54GL.

Durant el procés d'inici, i havent modificat el valor de la variable boot_wait (=on) de la nvram, l'equip es configura amb la direcció IP per defecte (192.168.1.1) durant aquest temps d'espera. Durant aquest període, i al rebre algun dels intents del servidor TFTP, s'inicia la transferència del fitxer i posterior flashejat de l'aparell. Cal esperar una estona i després, reiniciar l'equip. Per tal de tenir el WRT54GL totalment funcional en el cas d'utilitzar un sistema SquashFS és necessari tornar a reiniciar l'equip després de l'execució de l'script "firstboot".

A més, a l'hora de descarregar el fitxer amb el firmware es pot escollir entre tres versions diferents:

- **Micro:** És la versió mínima, amb base-files, base-files-brcm, bridge, busybox, dnsmasq, dropbear, ipkg-sh, iptables, iwlib, kmod-switch, kmod-brcm-wl, kmod-diag, kmod-wlcompat, mtd, nvram, uclibc, wificonf.
- **Bin:** Versió per defecte i escollida per aquest treball. Incorpora base-files, base-files-brcm, bridge, busybox, dnsmasq, dropbear, haserl, ipkg, iptables, iwlib, kmod-switch, kmod-brcm-wl, kmod-diag, kmod-ppp, kmod-pppoe, kmod-wlcompat, mtd, nvram, ppp, ppp-mod-pppoe, uclibc, webif, wificonf, wireless-tools.
- **Pptp:** Una versió amb menys funcionalitats, però amb suport per PPTP. Incorpora base-files, base-files-brcm, bridge, busybox, dnsmasq, dropbear, haserl, ipkg, iptables, iwlib, kmod-switch, kmod-brcm-wl, kmod-diag, kmod-ppp, kmod-gre, kmod-wlcompat, mtd, nvram, ppp, pptp, uclibc, webif, wificonf, wireless-tools.

A1.1.2. DD-WRT

DD-WRT és un projecte basat en el codi original de Linksys, Sveasoft Alchemy, també distribuït sota llicència GPL. En DD-WRT s'ha apostat per utilitzar tot el nucli de OpenWRT per al kernel. Incorpora una interfície web molt amigable i gran quantitat de software disponible.

La instal·lació de DD-WRT és igual de senzilla que en el cas de OpenWRT. Per a les actualitzacions de DD-WRT. Només comentar, que per a la migració de OpenWRT cap a DD-WRT és necessari seguir el procediment amb el servidor TFTP i el `boot_wait`.

La versió utilitzada és la v23, que incorpora les últimes novetats en quant a software disponible, entre els que cal nombrar el dimoni d'encaminament BIRD.

A1.1.3. Versions de Guifi.net

S'ha de comentar que les proves d'aquest estudi es realitzen amb les distribucions OpenWRT i DD-WRT. La primera no és compatible amb l'aplicació UnSolClick de Guifi.net. Aquesta funcionalitat no es vol utilitzar i el dimoni d'encaminament si és el mateix. Per tant, les configuracions resultants si són compatibles amb alguna de les distribucions de Guifi.net (Guifi.net/Drupal i www.comesfa.org/guifi.dir/DDWRT). Concretament, la DD-WRT si és compatible amb Guifi.net, encara que sota el nom DD-Guifi.

A1.2. Memòria persistent: NVRAM

La memòria persistent, o nvrn, conté la informació que utilitza el WRT54GL que s'engega per poder configurar el sistema adequadament. El seu contingut no s'esborra després que l'equip perdi l'alimentació elèctrica i és important conèixer el seu funcionament.

Per treballar amb la memòria nvrn del WRT54GL existeix una comanda amb el mateix nom, `nvrn`. Algunes de les instruccions que ofereix aquesta comanda són:

- Mostrar totes les variables: `nvrn show`. Mostra tot el llistat amb totes les variables del sistema.
- Mostrar el valor d'una variable: `nvrn get NomVariable`. Mostra el valor de la variable indicat en `NomVariable`.
- Modificar el valor d'una variable: `nvrn set NomVariable="Valor"`. Modifica el valor de la variable indicat en `NomVariable` amb el valor indicat a `Valor`. Les cometes són necessàries si dins el valor hi ha espais, sinó el sistema talla el valor al trobar el primer espai.

- Eliminar una variable: `nvramp unset NomVariable`. Elimina una variable del sistema.
- Acceptar tots els canvis realitzats: `nvramp commit`. Fins que no s'executa aquesta instrucció els canvis realitzats només són modificats en memòria RAM, per això si es treu la corrent a l'encaminador aquest recupera els valors anterior. Amb aquesta instrucció els valors modificats queden emmagatzemats permanentment.

Per obtenir els valors originals, després d'entrar en mode segur, és necessari executar la comanda “`mtid -r erase nvramp`”. El paràmetre “-r” reinicia l'equip. Abans d'executar la comanda és necessari modificar la variable `boot_wait` a `on`, sinó la nova partició `nvramp` s'escriu amb els valors de la cache del WRT54GL.

A1.3. Failsafe mode

Failsafe mode, o mode de recuperació, s'utilitza si el WRT54GL no s'inicia adequadament a causa d'un problema en un script d'inici, una configuració de tallafocs errònia, o qualsevol altre motiu semblant. Aquest mode d'operació funciona amb les imatges SquashFS.

Per accedir al mode segur és necessari polsar el botó de RESET durant dos segons abans que el gestor d'inici (boot loader) s'iniciï. En els models amb un LED etiquetat com DMZ s'ha de polsar el botó abans que s'il·lumini. En cas contrari, si polsa el botó de RESET durant el procés d'inici es regatejarà la memòria NVRAM als valors per defecte.

Iniciar en mode segur pot modificar l'adreça del dispositiu, invalidant les comunicacions. Per a solucionar-ho només és necessari esborrar la memòria cache de l'estació de treball utilitzada. Un cop es té accés al dispositiu es pot formatejar la partició JFFS2 (`/bin/firstboot`) o intentar corregir el problema (`/sbin/mount_root`). Si s'executa el primer havent muntat la partició JFFS2 amb la segona comanda no es formatejarà la partició, però si es substituiran els fitxers per enllaços simbòlics (el software es mantindrà però les modificacions als scripts no).

A1.4. El switch

Els Linksys WRT54GL utilitzen una interfície virtual anomenada pont que s'encarrega d'ajuntar dues o més interfícies reals com si pertanyessin a una sola subxarxa. En la configuració per defecte existeix un pont entre la interfície WLAN i els quatre ports de xarxa LAN.

El sistema disposa de l'aplicació `brctl` per crear, modificar i eliminar ponts dins el sistema. Els WRT54GL configuren el pont `br0` amb les interfícies `wlan0` i `eth1`, que equivalen a la LAN i el Wi-Fi.

Dins la nvram, la variable `lan_ifnames="vlan0 eth1 eth2 eth3"` indica que les interfícies anteriors formen part de la LAN del sistema, i a més apareix la variable `lan_ifname="br0"` indicant quina és la interfície que ha de rebre la configuració de la LAN. Així es pot modificar aquestes dues variables per modificar el comportament de la configuració del pont `br0`, i que no es creï al iniciar el sistema: `lan_ifnames=""` i `lan_ifname="vlan0"`.

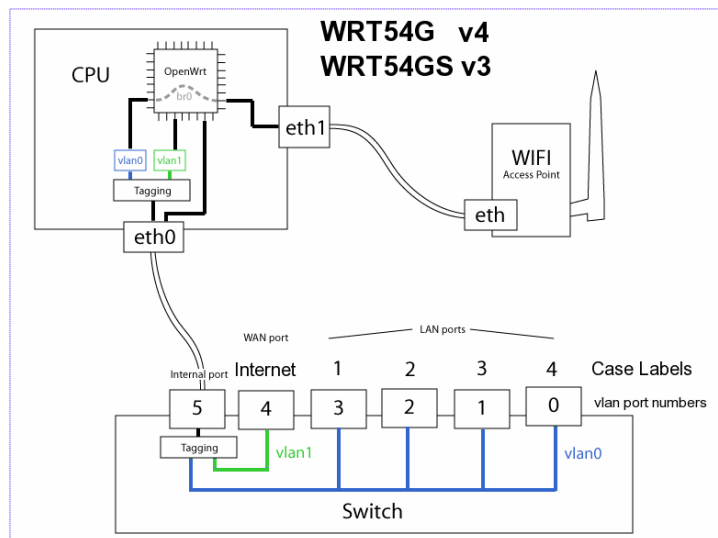


Fig. A1.1 Estructura interna del WR54GL

Les xarxes virtuals s'utilitzen per separar en diferents subxarxes els ports cablejats del WRT54GL. Gràcies a aquesta limitació, i en la forma que engega el linksys, es pot aconseguir que cadascun dels ports pertanyi a una subxarxa diferent.

Per defecte, el sistema crea quatre variables en la nvram que s'encarreguen de configurar les xarxes virtuals originals, que són dues.

```
vlan0hwname=eth0
vlan0ports="1 2 3 4 5"
```

```
vlan1hwname=eth0
vlan1ports="0 5"
```

Les variables "`vlanXhwname`" fan referència a la interfície on està assignada la xarxa virtual indicada. Les variables "`vlanXports`" fan referència als ports de la interfície que seran assignats a cada xarxa virtual. En aquests cas, la `vlan0` que pertany als ports 1 a 4 de la LAN, i la `vlan1` fa referència al port 0 d'Internet.

El valor 5 que es posa al final de les variables de ports fa referència a la pròpia interfície i no pot faltar en cap de les definicions, si s'omet aquest valor al final d'alguna de les variables dita VLAN deixarà de funcionar. Així mateix l'asterisc que es troba al final de la primera definició en la variable "`vlan0ports`" també és obligatòria per indicar quina VLAN té el control principal de la interfície. Si s'omet, cap dels ports funcionarà i les interfícies quedaran inservibles, i

difícilment podrà modificar-se de nou la variable per aconseguir que els ports de la interfície tornin a funcionar.

Per tant, per definir les xarxes virtuals necessàries per obtenir-ne una de diferent per cada port s'han de definir les següents variables, una per a cada VLAN:

```
vlan0hwname=et0
vlan1hwname=et0
vlan2hwname=et0
vlan3hwname=et0
vlan4hwname=et0

vlan0ports="0 5*"
vlan1ports="4 5"
vlan2ports="2 5"
vlan3ports="3 5"
vlan4ports="1 5"
```

La configuració que aquí es mostra assigna la vlan1 al port etiquetat com a WAN, la vlan3 al port 1, la vlan2 al port 2, la vlan4 al port 3 i la VLAN 0 al port 4 i és la que s'ha realitzat a tots els equips WRT54GL disponibles.

Un cop tenim les variables modificades i guardades en memòria amb la comanda "nvram commit", cada cop que es tregui l'alimentació i s'engegui el sistema tindrem a la nostra configuració la possibilitat de tenir fins 5 subxarxes diferents. S'ha de tenir en compte que a partir d'ara la configuració de la xarxa local que ve configurada per defecte en el Linksys, i que es pot modificar mitjançant la web només estarà assignada a els ports de la vlan0, tal com estava per defecte, així que s'ha de tenir clar com es configuren cadascuna de les xarxes i ports.

A1.5. Metodologia de treball

Un cop els WRT54GL tenen el sistema operatiu instal·lat i funcional, i a partir de la nvram es pot modificar la configuració del switch l'equip es pot veure com un ordinador amb Linux

Cal tenir en compte alguna consideració especial a l'escriptura a disc si s'utilitza la opció SquashFS en el cas d'OpenWRT (en DD-WRT no). Per editar fitxers és necessari esborrar la còpia actual i obtenir el nou fitxer. Un exemple d'ús és el següent:

```
rm /etc/ipkg.conf
cp /rom/etc/ipkg.conf /etc/ipkg.conf
vim /etc/ipkg.conf
```

Per realitzar una còpia de la instal·lació actual, i obtenir un fitxer cal executar la comanda "dd if=/dev/mtdblock/1 of=/tmp/firmware.trx" i per restablir aquesta imatge "mtd -r write firmware.trx linux". El paràmetre "-r" reiniciarà l'equip un cop s'hagi finalitzat l'escriptura. De fet, aquest és un procediment alternatiu a l'ús del servidor TFTP, comentat a l'inici d'aquest capítol.

A1.6. Bird

BIRD és una implementació d'un encaminador IP que executa diferents protocols. Ha estat desenvolupat a la Facultat de Matemàtiques i Física, Charles University, Pagra, República Txeca com un projecte d'estudis. Aquest pot ser lliurement distribuït sota els termes de la Llicència Pública General GNU.

Les característiques de BIRD són:

- Ambdós protocols, IPv4 i IPv6.
- Múltiples taules d'encaminament.
- Border Gateway Protocol (BGPv4).
- Routing Information Protocol (RIPv2).
- Open Shortest Path First protocol (OSPFv2).
- Un protocol virtual per intercanvi de rutes entre diferents taules d'encaminament en un únic client.
- Una interfície de línia de comandes que permet el control en - línia i la inspecció de l'estat del dimoni.
- Reconfiguració per programari (no es necessari l'ús de complexes comandes en-línia per canvia la configuració, només cal editar el fitxer de configuració i notificar a BIRD que el rellegeixi i suaument canviï a la nova configuració, no distribuint els protocols d'encaminament excepte que siguin afectats pels canvis de la configuració).
- un potent llenguatge per filtrat de rutes.

BIRD te una o més taules d'encaminament, les quals poden o no poden ser sincronitzades amb el kernel del Sistema Operatiu i que poden o no poden ser sincronitzats amb cada altre (mirar el protocol Pipe). Cada taula d'encaminament conté una llista de les rutes conegudes. Cada ruta consisteix de:

- El prefix de la xarxa d'aquesta ruta és per (adreça de xarxa i longitud del prefix -- el nombre de bits que forma la part de l'adreça de la xarxa; també coneguda com a màscara de xarxa).
- preferència d'aquesta ruta.
- adreça IP de l'encaminador del qual em obtingut la ruta.
- adreça IP de l'encaminador on hem de reenviar els paquets per utilitzar aquesta ruta.
- altres atributs comuns a totes les rutes
- atributs dinàmics definits pels protocols que poden o no poden estar presents (normalment mètriques de protocol).

La taula d'encaminament manté múltiples entrades per una xarxa, però al menys una entrada per una xarxa i un protocol. La entrada amb la preferència més alta s'utilitza per l'encaminament. Si hi ha més entrades amb les mateixes preferències i són del mateix protocol, el protocol decideix (normalment d'acord amb les mètriques). Si no ho son, una ordre interna s'utilitza per trencar

l'empat. Es pot obtenir la llista dels atributs de la ruta en la secció d'atributs d'Encaminament.

Cada protocol està connectat a una taula d'encaminament a través de dos filtres, els quals poden acceptar, refusar i modificar les rutes. Un filtre *export* comprova les rutes passades des de la taula d'encaminament al protocol, un filtre *import* comprova les rutes en la direcció oposada. Quan les taules d'encaminament obtenen una ruta des d'un protocol, aquesta recalcula la ruta seleccionada i els broadcasts de tots els protocols connectats a la taula. Els protocols normalment envien la modificació a altres encaminadors dins la xarxa.

Tota la configuració de Bird es realitza per fitxers, sintaxi dels quals es pot trobar a la documentació on-line o als manuals del propi software.

A1.7. Quagga

Quagga és un altre dimoni d'encaminament amb les mateixes funcionalitats que l'anterior. Per tant, aporta encaminament RIPv1, RIPv2, RIPv6, OSPFv2, OSPFv3, BGP-4 y BGP-4+, entre d'altres. Com a factor diferencial, es compatible amb la funcionalitat de BGP Route Reflector y Route Server, d'especial interès al llarg d'aquest estudi.

Quagga utilitza una arquitectura de software avançada i pensada per proporcionar gran qualitat. Compta amb un motor multiservidor d'encaminament i una interfície d'accés unificada per a tots els protocols. El software tradicional d'encaminament està compost per un programa o procés únic que proporciona totes les funcionalitats. Quagga està format per un gestor de routing (zebra) i una sèrie de programes específics que executen cadascun del protocols. Aquesta arquitectura presentada i la visió multiprocés permeten un sistema més fàcilment extensible i totalment modular.

Per acabar, comentar que aquest software és addicional i per tant s'ha d'instal·lar sota la distribució OpenWRT (DD-WRT ja compta a Bird per defecte). Per a fer-ho, només és necessari utilitzar l'aplicació `ipkg`. De fet, aquesta és la eina utilitzada per a instal·lar qualsevol software addicional en OpenWRT.

```
ipkg update
ipkg install quagga
ipkg install quagga-ospfd
ipkg install quagga-bgpd
```

ANNEX 2. CONFIGURACIONS OSPF

A2.1. Introducció

L'objectiu d'aquest annex és recollir les configuracions comentades i el màxim d'informació OSPF que es pot obtenir a partir de la consola d'administració. Tota la informació aquí recollida ha de facilitar al màxim la reproduïbilitat dels escenaris.

A2.2. Escenari 2

En aquest segon escenari d'OSPF es volen testejar algunes de les funcionalitats avançades del protocol i la creació d'enllaços virtuals:

- Particionat de la xarxa OSPF en varies àrees. Experiència amb assignació d'IPs.
- Tolerància a fallides.
- Enllaços virtuals.

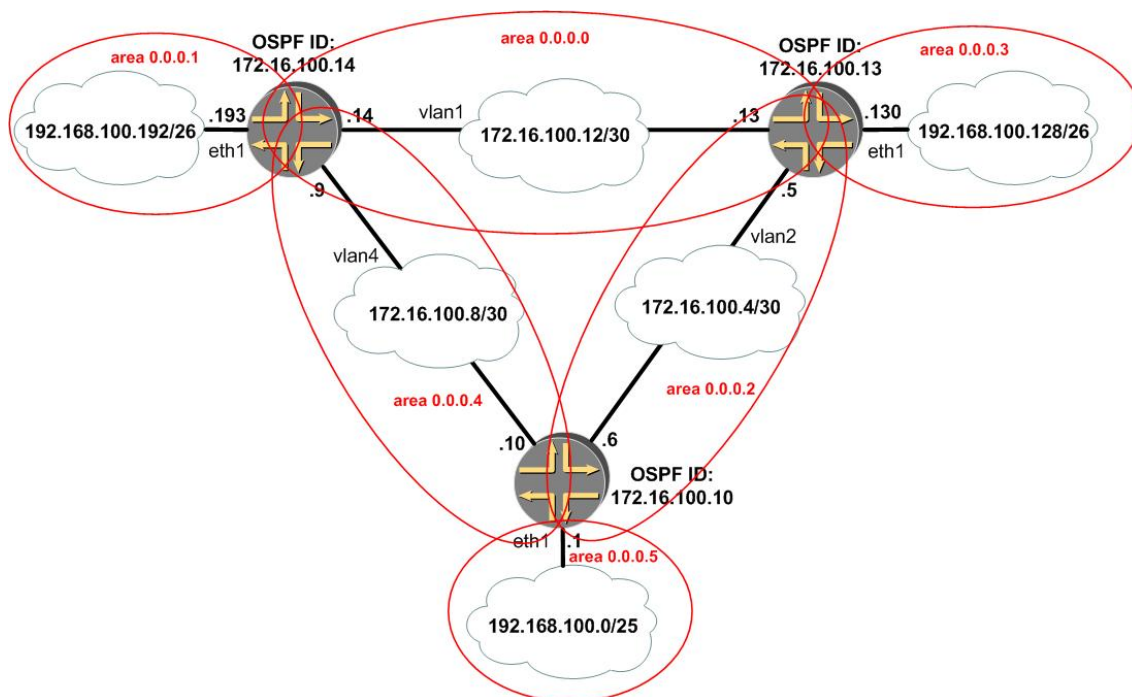


Fig. A2.1 Segon escenari OSPF

A2.2.1. Router 4

La configuració final de l'encaminador ABR de les àrees 0.0.0.0, 0.0.0.1 i 0.0.0.4 queda de la següent forma:

```
Current configuration:
!
hostname OpenWRT@AP4
log stdout
!
password ap4
enable password ap4
!
interface eth0
!
interface eth1
 ip address 192.168.100.193/26
!
interface lo
 ip address 127.0.0.1/8
!
interface vlan0
 ip address 192.168.1.1/24
!
interface vlan1
 ip address 172.16.100.14/30
 ip ospf cost 10
!
interface vlan2
!
interface vlan3
!
interface vlan4
 ip address 172.16.100.9/30
 ip ospf cost 10
!
router ospf
 ospf router-id 172.16.100.14
 redistribute kernel metric 1
 redistribute connected metric 1
 network 172.16.100.8/30 area 0.0.0.4
 network 172.16.100.12/30 area 0.0.0.0
 network 192.168.100.192/26 area 0.0.0.1
 area 0.0.0.1 stub
 area 0.0.0.4 virtual-link 172.16.100.10
!
ip forwarding
ipv6 forwarding
!
line vty
!
```

El procés de veïnatge funciona perfectament també aplicat en enllaços virtuals:

```
OpenWrt# show ip ospf neighbor
```

Neighbor	ID	Pri	State	Dead Time	Address	Interface
172.16.100.13		1	Full/Backup	00:00:35	172.16.100.13	vlan1:172.16.100.14
0	0	0				
172.16.100.10		1	Full/Backup	00:00:32	172.16.100.10	vlan4:172.16.100.9
0	0	0				
172.16.100.10		1	Full/DROther	00:00:39	172.16.100.10	VLINK0
0	0					0

En aquesta traça ja es pot veure que el procés de veïnatge (també anomenat d'adjacència) pot atorgar diferents rols a les entitats OSPF. El següent diagrama mostra el procés complet:

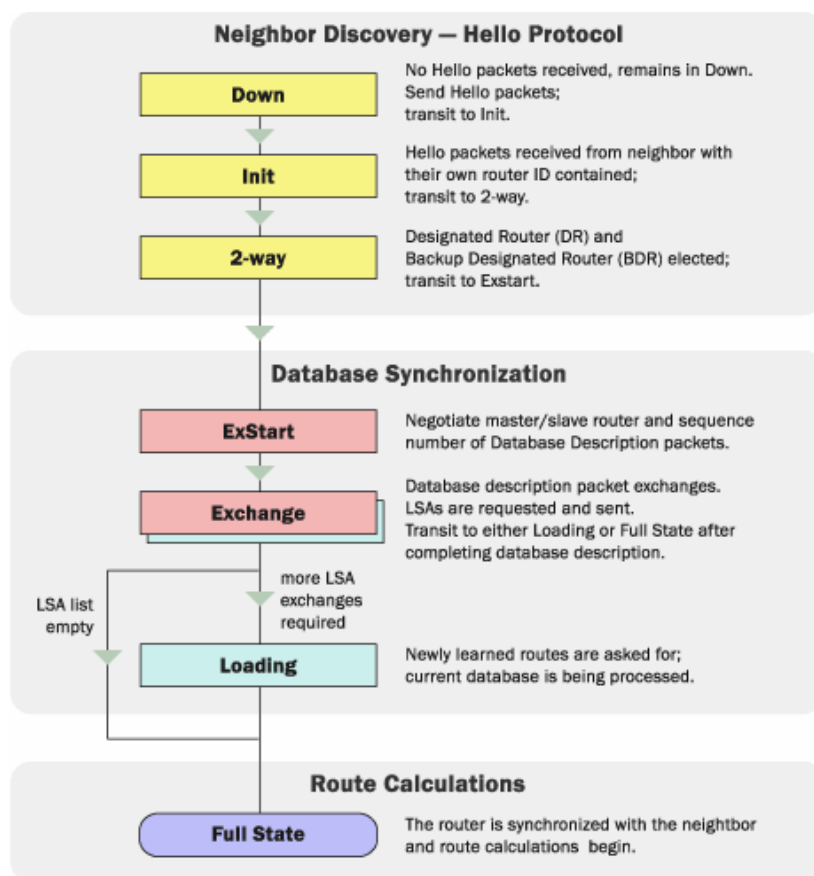


Fig. A2.2 Procés d'adjacència OSPF

I les rutes completes de la taula d'encaminament:

```
OpenWrt# show ip route
Codes: K - kernel route, C - connected, S - static, R - RIP, O - OSPF,
       I - ISIS, B - BGP, > - selected route, * - FIB route

C>* 127.0.0.0/8 is directly connected, lo
O>* 172.16.100.4/30 [110/20] via 172.16.100.13, vlan1, 00:19:19
                        via 172.16.100.10, vlan4, 00:19:19
O 172.16.100.8/30 [110/10] is directly connected, vlan4, 3d21h16m
C>* 172.16.100.8/30 is directly connected, vlan4
O 172.16.100.12/30 [110/10] is directly connected, vlan1, 3d21h16m
C>* 172.16.100.12/30 is directly connected, vlan1
O>* 192.168.100.0/25 [110/20] via 172.16.100.10, vlan4, 00:19:17
O>* 192.168.100.128/26 [110/20] via 172.16.100.13, vlan1, 1d22h43m
O 192.168.100.192/26 [110/10] is directly connected, eth1, 3d03h39m
C>* 192.168.100.192/26 is directly connected, eth1
```

La base de dades OSPF és extensa, ja que l'encaminador n'ha de mantenir una de diferent per a cada àrea. En aquesta traça ja es veuen les tres juntes.

```
OpenWrt# show ip ospf database
```


OSPF Router with ID (172.16.100.14)

Router Link States (Area 0.0.0.0)

Link ID	ADV Router	Age	Seq#	CkSum	Link count
172.16.100.10	172.16.100.10	1214	0x8000000c	0xa96f	2
172.16.100.13	172.16.100.13	1216	0x80000063	0x3e80	2
172.16.100.14	172.16.100.14	1214	0x800000c9	0xd779	2

Net Link States (Area 0.0.0.0)

Link ID	ADV Router	Age	Seq#	CkSum
172.16.100.14	172.16.100.14	983	0x8000005e	0xf94a

Summary Link States (Area 0.0.0.0)

Link ID	ADV Router	Age	Seq#	CkSum	Route
172.16.100.4	172.16.100.10	1304	0x80000001	0x32d3	172.16.100.4/30
172.16.100.4	172.16.100.13	1143	0x8000005e	0x6540	172.16.100.4/30
172.16.100.8	172.16.100.10	1329	0x80000001	0x0af7	172.16.100.8/30
172.16.100.8	172.16.100.14	551	0x800000ba	0x7ec5	172.16.100.8/30
192.168.100.0	172.16.100.10	317	0x80000002	0x4197	192.168.100.0/25
192.168.100.128	172.16.100.13	1443	0x8000005d	0xf4c4	192.168.100.128/26
192.168.100.192	172.16.100.14	1452	0x80000097	0xf746	192.168.100.192/26

ASBR-Summary Link States (Area 0.0.0.0)

Link ID	ADV Router	Age	Seq#	CkSum
172.16.100.10	172.16.100.13	1221	0x80000001	0xe711
172.16.100.10	172.16.100.14	1220	0x80000001	0xe116
172.16.100.13	172.16.100.10	1239	0x80000001	0xdb1d
172.16.100.14	172.16.100.10	237	0x80000003	0xcd28

Router Link States (Area 0.0.0.1 [Stub])

Link ID	ADV Router	Age	Seq#	CkSum	Link count
172.16.100.14	172.16.100.14	565	0x800000c0	0x3d9f	1

Summary Link States (Area 0.0.0.1 [Stub])

Link ID	ADV Router	Age	Seq#	CkSum	Route
0.0.0.0	172.16.100.14	1141	0x80000097	0xc0d7	0.0.0.0/0
172.16.100.4	172.16.100.14	1822	0x8000005d	0xe3b9	172.16.100.4/30
172.16.100.8	172.16.100.14	751	0x8000005e	0x554d	172.16.100.8/30
172.16.100.12	172.16.100.14	1392	0x8000005d	0x2f70	172.16.100.12/30
192.168.100.0	172.16.100.14	1213	0x80000001	0xad20	192.168.100.0/25
192.168.100.128	172.16.100.14	891	0x8000005e	0x6f40	192.168.100.128/26

Router Link States (Area 0.0.0.4)

Link ID	ADV Router	Age	Seq#	CkSum	Link count
172.16.100.10	172.16.100.10	1215	0x8000000f	0x8efa	1
172.16.100.14	172.16.100.14	1216	0x800000d4	0xb408	1

Net Link States (Area 0.0.0.4)

Link ID	ADV Router	Age	Seq#	CkSum
172.16.100.9	172.16.100.14	1225	0x80000001	0xb8f0

Summary Link States (Area 0.0.0.4)

Link ID	ADV Router	Age	Seq#	CkSum	Route
172.16.100.4	172.16.100.10	1305	0x80000001	0x32d3	172.16.100.4/30
172.16.100.12	172.16.100.14	471	0x800000ba	0x56e9	172.16.100.12/30
192.168.100.0	172.16.100.10	1600	0x80000002	0x4197	192.168.100.0/25
192.168.100.128	172.16.100.10	277	0x80000002	0x22eb	192.168.100.128/26
192.168.100.128	172.16.100.14	1231	0x8000005d	0x535b	192.168.100.128/26
192.168.100.192	172.16.100.14	1311	0x80000097	0xf746	192.168.100.192/26

ASBR-Summary Link States (Area 0.0.0.4)

Link ID	ADV Router	Age	Seq#	CkSum
172.16.100.13	172.16.100.10	1221	0x80000001	0xdb1d
172.16.100.13	172.16.100.14	380	0x8000005e	0x098e

Per últim, un parell de mesures sobre els consums de CPU i memòria en regim permanent i les rutes del kernel del sistema:

Mem: 11072K used, 3304K free, 0K shrd, 524K buff, 3588K cached
 Load average: 0.00, 0.00, 0.00 (State: S=sleeping R=running, W=waiting)

PID	USER	STATUS	RSS	PPID	%CPU	%MEM	COMMAND
2840	root	R	408	2749	1.9	2.8	top
2748	root	S	604	511	0.9	4.2	dropbear
2073	quagga	S	1696	1	0.0	11.7	bgpd
2061	quagga	S	1172	1	0.0	8.1	ospfd
2049	quagga	S	652	1	0.0	4.5	zebra

A2.2.2. Router 2

La configuració funcional de l'ABR de les àrees 0.0.0.0, 0.0.0.2 i 0.0.0.3 queda de la següent forma:

```
Current configuration:
!
hostname OpenWRT@AP2
log stdout
!
interface eth0
!
interface eth1
 ip address 192.168.100.129/26
 ip ospf cost 10
!
interface lo
 ip address 127.0.0.1/8
!
interface vlan0
!
interface vlan1
 ip address 172.16.100.13/30
 ip ospf cost 10
!
interface vlan2
 ip address 172.16.100.5/30
 ip ospf cost 10
!
interface vlan3
!
interface vlan4
!
router ospf
 ospf router-id 172.16.100.13
 redistribute kernel metric 1
 redistribute connected metric 1
 network 172.16.100.4/30 area 0.0.0.2
 network 172.16.100.12/30 area 0.0.0.0
 network 192.168.100.128/26 area 0.0.0.3
 area 0.0.0.3 stub no-summary
 area 0.0.0.2 virtual-link 172.16.100.10
!
```

```

ip forwarding
ipv6 forwarding
!
line vty
!

```

Les taules d'encaminament i la base de dades amb la topologia de la xarxa queden:

```

OpenWrt# show ip ospf route
===== OSPF network routing table =====
N    172.16.100.4/30      [10] area: 0.0.0.2
      directly attached to vlan2
N IA 172.16.100.8/30      [20] area: 0.0.0.0
      via 172.16.100.14, vlan1
      via 172.16.100.6, vlan2
N    172.16.100.12/30     [10] area: 0.0.0.0
      directly attached to vlan1
N IA 192.168.100.0/25     [20] area: 0.0.0.0
      via 172.16.100.6, vlan2
N    192.168.100.128/26  [10] area: 0.0.0.3
      directly attached to eth1
N IA 192.168.100.192/26 [20] area: 0.0.0.0
      via 172.16.100.14, vlan1

===== OSPF router routing table =====
R    172.16.100.10        [10] area: 0.0.0.2, ABR, ASBR
      via 172.16.100.6, vlan2
R    172.16.100.14        [10] area: 0.0.0.0, ABR, ASBR
      via 172.16.100.14, vlan1

```

```
OpenWrt# show ip ospf database
```

```
OSPF Router with ID (172.16.100.13)
```

```
Router Link States (Area 0.0.0.0)
```

Link ID	ADV Router	Age	Seq#	CkSum	Link count
172.16.100.10	172.16.100.10	1028	0x8000000c	0xa96f	2
172.16.100.13	172.16.100.13	1029	0x80000063	0x3e80	2
172.16.100.14	172.16.100.14	1029	0x800000c9	0xd779	2

```
Net Link States (Area 0.0.0.0)
```

Link ID	ADV Router	Age	Seq#	CkSum
172.16.100.14	172.16.100.14	798	0x8000005e	0xf94a

```
Summary Link States (Area 0.0.0.0)
```

Link ID	ADV Router	Age	Seq#	CkSum	Route
172.16.100.4	172.16.100.10	1120	0x80000001	0x32d3	172.16.100.4/30
172.16.100.4	172.16.100.13	956	0x8000005e	0x6540	172.16.100.4/30
172.16.100.8	172.16.100.10	1145	0x80000001	0x0af7	172.16.100.8/30
172.16.100.8	172.16.100.14	366	0x800000ba	0x7ec5	172.16.100.8/30
192.168.100.0	172.16.100.10	132	0x80000002	0x4197	192.168.100.0/25
192.168.100.128	172.16.100.13	1256	0x8000005d	0xf4c4	192.168.100.128/26
192.168.100.192	172.16.100.14	1267	0x80000097	0xf746	192.168.100.192/26

```
ASBR-Summary Link States (Area 0.0.0.0)
```

Link ID	ADV Router	Age	Seq#	CkSum
172.16.100.10	172.16.100.13	1034	0x80000001	0xe711
172.16.100.10	172.16.100.14	1035	0x80000001	0xe116
172.16.100.13	172.16.100.10	1054	0x80000001	0xdb1d
172.16.100.14	172.16.100.10	52	0x80000003	0xcd28

Router Link States (Area 0.0.0.2)

Link ID	ADV Router	Age	Seq#	CkSum	Link count
172.16.100.10	172.16.100.10	1030	0x8000000a	0x385e	1
172.16.100.13	172.16.100.13	1030	0x80000066	0x41f3	1

Net Link States (Area 0.0.0.2)

Link ID	ADV Router	Age	Seq#	CkSum
172.16.100.5	172.16.100.13	1039	0x80000001	0xdcd2

Summary Link States (Area 0.0.0.2)

Link ID	ADV Router	Age	Seq#	CkSum	Route
172.16.100.8	172.16.100.10	1144	0x80000001	0x0af7	172.16.100.8/30
172.16.100.12	172.16.100.13	536	0x8000005e	0x1588	172.16.100.12/30
192.168.100.0	172.16.100.10	1565	0x80000002	0x4197	192.168.100.0/25
192.168.100.128	172.16.100.13	806	0x8000005d	0xf4c4	192.168.100.128/26
192.168.100.192	172.16.100.10	1029	0x80000001	0xa12d	192.168.100.192/26
192.168.100.192	172.16.100.13	1787	0x8000005d	0xd698	192.168.100.192/26

ASBR-Summary Link States (Area 0.0.0.2)

Link ID	ADV Router	Age	Seq#	CkSum
172.16.100.14	172.16.100.10	1035	0x80000001	0xd126
172.16.100.14	172.16.100.13	1527	0x8000005d	0x0791

Router Link States (Area 0.0.0.3 [Stub])

Link ID	ADV Router	Age	Seq#	CkSum	Link count
172.16.100.13	172.16.100.13	640	0x80000060	0x8bf3	1

Summary Link States (Area 0.0.0.3 [Stub])

Link ID	ADV Router	Age	Seq#	CkSum	Route
0.0.0.0	172.16.100.13	486	0x8000005e	0x3999	0.0.0.0/0

Igual que abans, també es realitzen un parell de captures dels consums de CPU i memòria. Es comprova que son semblants a l'anterior:

```
Mem: 9172K used, 5204K free, 0K shrd, 488K buff, 3440K cached
Load average: 0.00, 0.00, 0.00 (State: S=sleeping R=running, W=waiting)
```

```

PID USER      STATUS  RSS  PPID  %CPU  %MEM  COMMAND
1894 root        R       404  1870   2.9   2.8  top
1419 quagga     S      1384    1   0.0   9.6  ospfd
1279 quagga     S       684    1   0.0   4.7  zebra
```

A2.2.3. Router 3

Finalment, la configuració del tercer encaminador és la següent. Igual que els anterior, també pren el rol d'ABR, però aquest cop per les arees 0.0.0.2, 0.0.0.4 i 0.0.0.5:

```
hostname OpenWRT@AP3
log stdout
!
debug ospf event
debug ospf packet all
!
```

```

interface eth0
!
interface eth1
 ip address 192.168.100.1/25
 ip ospf cost 10
!
interface lo
 ip address 127.0.0.1/8
!
interface vlan0
 ip address 192.168.1.1/24
!
interface vlan1
!
interface vlan2
 ip address 172.16.100.6/30
 ip ospf cost 10
!
interface vlan3
!
interface vlan4
 ip address 172.16.100.10/30
 ip ospf cost 10
!
router ospf
 ospf router-id 172.16.100.10
 redistribute kernel metric 1
 redistribute connected metric 1
 network 172.16.100.4/30 area 0.0.0.2
 network 172.16.100.8/30 area 0.0.0.4
 network 192.168.100.1/25 area 0.0.0.5
 area 0.0.0.2 virtual-link 172.16.100.13
 area 0.0.0.4 virtual-link 172.16.100.14
!
ip forwarding
ipv6 forwarding
!
line vty
!

```

Amb els corresponents veïns:

```
OpenWrt# show ip ospf neighbor
```

Neighbor ID	Pri	State	Dead Time	Address	Interface
172.16.100.13	1	Full/DROther	00:00:31	172.16.100.5	VLINK0
0 0 0					
172.16.100.14	1	Full/DROther	00:00:30	172.16.100.9	VLINK1
0 0 0					
172.16.100.13	1	Full/DR		00:00:33	172.16.100.5
vlan2:172.16.100.6	0	0 0			
172.16.100.14	1	Full/DR		00:00:33	172.16.100.9
vlan4:172.16.100.10	0	0 0			

A partir de la traça anterior es pot veure que s'han configurat dos enllaços virtuals entre l'àrea 0.0.0.5 i la 0.0.0.0. Les taules d'encaminament resultants són:

```

OpenWrt# show ip ospf route
===== OSPF network routing table =====
N    172.16.100.4/30    [10] area: 0.0.0.2
      directly attached to vlan2
N    172.16.100.8/30    [10] area: 0.0.0.4
      directly attached to vlan4
N    172.16.100.12/30   [20] area: 0.0.0.0

```

```

                via 172.16.100.5, vlan2
                via 172.16.100.9, vlan4
N    192.168.100.0/25    [10] area: 0.0.0.5
                directly attached to eth1
N IA 192.168.100.128/26 [20] area: 0.0.0.0
                via 172.16.100.5, vlan2
N IA 192.168.100.192/26 [20] area: 0.0.0.0
                via 172.16.100.9, vlan4

===== OSPF router routing table =====
R    172.16.100.13    [10] area: 0.0.0.2, ABR, ASBR
                via 172.16.100.5, vlan2
R    172.16.100.14    [10] area: 0.0.0.4, ABR, ASBR
                via 172.16.100.9, vlan4

```

La base de dades d'encaminament és la més gran de les tres:

OSPF Router with ID (172.16.100.10)

Router Link States (Area 0.0.0.0)

Link ID	ADV Router	Age	Seq#	CkSum	Link count
172.16.100.10	172.16.100.10	619	0x8000000c	0xa96f	2
172.16.100.13	172.16.100.13	622	0x80000063	0x3e80	2
172.16.100.14	172.16.100.14	621	0x800000c9	0xd779	2

Net Link States (Area 0.0.0.0)

Link ID	ADV Router	Age	Seq#	CkSum
172.16.100.14	172.16.100.14	390	0x8000005e	0xf94a

Summary Link States (Area 0.0.0.0)

Link ID	ADV Router	Age	Seq#	CkSum	Route
172.16.100.4	172.16.100.10	631	0x80000001	0x32d3	172.16.100.4/30
172.16.100.4	172.16.100.13	549	0x8000005e	0x6540	172.16.100.4/30
172.16.100.8	172.16.100.10	631	0x80000001	0x0af7	172.16.100.8/30
172.16.100.8	172.16.100.14	1761	0x800000b9	0x80c4	172.16.100.8/30
192.168.100.0	172.16.100.10	631	0x80000001	0x4396	192.168.100.0/25
192.168.100.128	172.16.100.13	850	0x8000005d	0xf4c4	192.168.100.128/26
192.168.100.192	172.16.100.14	860	0x80000097	0xf746	192.168.100.192/26

ASBR-Summary Link States (Area 0.0.0.0)

Link ID	ADV Router	Age	Seq#	CkSum
172.16.100.10	172.16.100.13	628	0x80000001	0xe711
172.16.100.10	172.16.100.13	628	0x80000001	0xe711
172.16.100.10	172.16.100.14	628	0x80000001	0xe116
172.16.100.13	172.16.100.10	627	0x80000001	0xdb1d
172.16.100.14	172.16.100.10	1438	0x80000002	0xcf27

Router Link States (Area 0.0.0.2)

Link ID	ADV Router	Age	Seq#	CkSum	Link count
172.16.100.10	172.16.100.10	621	0x8000000a	0x385e	1
172.16.100.13	172.16.100.13	624	0x80000066	0x41f3	1

Net Link States (Area 0.0.0.2)

Link ID	ADV Router	Age	Seq#	CkSum
172.16.100.5	172.16.100.13	633	0x80000001	0xdcd2

Summary Link States (Area 0.0.0.2)

Link ID	ADV Router	Age	Seq#	CkSum	Route
172.16.100.8	172.16.100.10	737	0x80000001	0x0af7	172.16.100.8/30
172.16.100.12	172.16.100.13	129	0x8000005e	0x1588	172.16.100.12/30

```

192.168.100.0 172.16.100.10 1158 0x80000002 0x4197 192.168.100.0/25
192.168.100.128 172.16.100.13 399 0x8000005d 0xf4c4 192.168.100.128/26
192.168.100.192 172.16.100.10 620 0x80000001 0xa12d 192.168.100.192/26
192.168.100.192 172.16.100.13 1380 0x8000005d 0xd698 192.168.100.192/26

```

ASBR-Summary Link States (Area 0.0.0.2)

```

Link ID      ADV Router    Age  Seq#      CkSum
172.16.100.14 172.16.100.10 627 0x80000001 0xd126
172.16.100.14 172.16.100.13 1120 0x8000005d 0x0791

```

Router Link States (Area 0.0.0.4)

```

Link ID      ADV Router    Age  Seq#      CkSum  Link count
172.16.100.10 172.16.100.10 620 0x8000000f 0x8efa 1
172.16.100.14 172.16.100.14 623 0x800000d4 0xb408 1

```

Net Link States (Area 0.0.0.4)

```

Link ID      ADV Router    Age  Seq#      CkSum
172.16.100.9 172.16.100.14 632 0x80000001 0xb8f0

```

Summary Link States (Area 0.0.0.4)

```

Link ID      ADV Router    Age  Seq#      CkSum  Route
172.16.100.4 172.16.100.10 712 0x80000001 0x32d3 172.16.100.4/30
172.16.100.12 172.16.100.14 1719 0x800000b9 0x58e8 172.16.100.12/30
192.168.100.0 172.16.100.10 1007 0x80000002 0x4197 192.168.100.0/25
192.168.100.128 172.16.100.10 620 0x80000001 0x24ea 192.168.100.128/26
192.168.100.128 172.16.100.14 638 0x8000005d 0x535b 192.168.100.128/26
192.168.100.192 172.16.100.14 718 0x80000097 0xf746 192.168.100.192/26

```

ASBR-Summary Link States (Area 0.0.0.4)

```

Link ID      ADV Router    Age  Seq#      CkSum
172.16.100.13 172.16.100.10 627 0x80000001 0xdb1d
172.16.100.13 172.16.100.14 1579 0x8000005d 0x0b8d

```

Router Link States (Area 0.0.0.5)

```

Link ID      ADV Router    Age  Seq#      CkSum  Link count
172.16.100.10 172.16.100.10 594 0x80000002 0xd7c8 1

```

Summary Link States (Area 0.0.0.5)

```

Link ID      ADV Router    Age  Seq#      CkSum  Route
172.16.100.4 172.16.100.10 631 0x80000001 0x32d3 172.16.100.4/30
172.16.100.8 172.16.100.10 631 0x80000001 0x0af7 172.16.100.8/30
172.16.100.12 172.16.100.10 620 0x80000001 0x46ad 172.16.100.12/30
192.168.100.128 172.16.100.10 620 0x80000001 0x24ea 192.168.100.128/26
192.168.100.192 172.16.100.10 620 0x80000001 0xa12d 192.168.100.192/26

```

ASBR-Summary Link States (Area 0.0.0.5)

```

Link ID      ADV Router    Age  Seq#      CkSum
172.16.100.13 172.16.100.10 627 0x80000001 0xdb1d
172.16.100.14 172.16.100.10 627 0x80000001 0xd126

```

I per acabar, les dues captues de rendiment de l'encaminador en regim permanent:

```

Mem: 11080K used, 3296K free, 0K shrd, 828K buff, 4528K cached
Load average: 0.00, 0.00, 0.00 (State: S=sleeping R=running, W=waiting)

```

```

PID USER      STATUS  RSS  PPID  %CPU  %MEM  COMMAND
9386 root       S       648   513   0.7   4.5  dropbear
9415 root       R       412  9387   0.7   2.8  top

```

```

9362 quagga s 1412 1 0.1 9.8 ospfd
9350 quagga s 840 1 0.0 5.8 zebra

```

A2.3. Escenari 3

El tercer escenari pretén esser una configuració d'exemple per a Guifi.net del que pot esser per exemple la xarxa d'un municipi o d'una comarca, ja que s'utilitzen diverses xarxes d'usuari i un particionat en àrees força extensiu. Algun dels objectius en aquest escenari són:

- Veure el funcionament del procés de sumarís en configuracions trivials.
- Veure el funcionament del procés de sumarís en configuracions no tant trivials.
- Comprovar de nou la tolerància a fallides, aquest cop amb una area 0.0.0.0 més extensa.

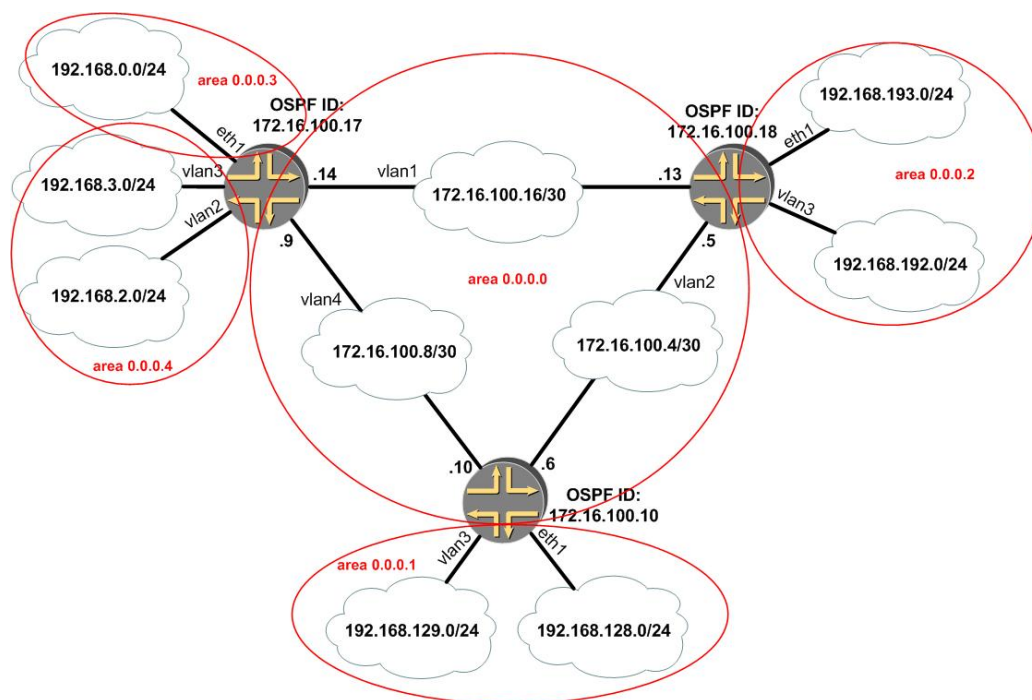


Fig. A2.3 Tercer escenari OSPF

A2.3.1. Router 4

La configuració de l'encaminador 4, amb OSPF ID 172.16.100.17 és la següent. Notar que en aquest punt ja es configura la sumarització d'adreces entre àrees:

```

hostname OpenWRT@AP4
log stdout
!
password ap4
enable password ap4
!

```



```

N    172.16.100.8/30      [10] area: 0.0.0.0
                        directly attached to vlan4
N    172.16.100.16/30   [10] area: 0.0.0.0
                        directly attached to vlan1
N    192.168.0.0/24     [10] area: 0.0.0.3
                        directly attached to eth1
N    192.168.2.0/24     [10] area: 0.0.0.4
                        directly attached to vlan2
N    192.168.3.0/24     [10] area: 0.0.0.4
                        directly attached to vlan3
N IA 192.168.128.0/18  [20] area: 0.0.0.0
                        via 172.16.100.10, vlan4
N IA 192.168.192.0/18  [20] area: 0.0.0.0
                        via 172.16.100.18, vlan1

===== OSPF router routing table =====
R    172.16.100.10      [10] area: 0.0.0.0, ABR
                        via 172.16.100.10, vlan4
R    172.16.100.18     [10] area: 0.0.0.0, ABR
                        via 172.16.100.18, vlan1

```

S'aprecia un consum de recursos lleugerament inferior a l'escenari anterior:

Mem: 10260K used, 4116K free, 0K shrd, 548K buff, 3752K cached
 Load average: 0.00, 0.00, 0.00 (State: S=sleeping R=running, W=waiting)

PID	USER	STATUS	RSS	PPID	%CPU	%MEM	COMMAND
4760	root	R	408	4717	3.8	2.8	top
4660	quagga	S	1368	1	0.0	9.5	ospfd
4648	quagga	S	860	1	0.0	5.9	zebra

A2.3.2. Router 3

En aquest cas, la configuració de l'encaminador ABR de l'àrea 0.0.0.1 és la següent:

```

!
hostname OpenWRT@AP3
log stdout
!
debug ospf event
debug ospf packet all
!
interface eth0
!
interface eth1
 ip address 192.168.128.1/24
 ip ospf cost 10
!
interface lo
 ip address 127.0.0.1/8
!
interface vlan0
 ip address 192.168.1.1/24
!
interface vlan1
!
interface vlan2
 ip address 172.16.100.6/30
 ip ospf cost 10
!
interface vlan3
 ip address 192.168.129.1/24

```

```

ip ospf cost 10
!
interface vlan4
ip address 172.16.100.10/30
ip ospf cost 10
!
router ospf
ospf router-id 172.16.100.10
network 172.16.100.4/30 area 0.0.0.0
network 172.16.100.8/30 area 0.0.0.0
network 172.16.100.16/30 area 0.0.0.0
network 192.168.128.0/24 area 0.0.0.1
network 192.168.129.0/24 area 0.0.0.1
area 0.0.0.1 stub no-summary
area 0.0.0.1 range 192.168.128.0/24 substitute 192.168.128.0/18
area 0.0.0.1 range 192.168.129.0/24 not-advertise
!
ip forwarding
ipv6 forwarding
!
line vty
!

```

I les rutes, també agregades, es poden veure per mitjà de les comandes show ip route, en les seves diferents variants:

```

OpenWrt# show ip ospf route
===== OSPF network routing table =====
N    172.16.100.4/30      [10] area: 0.0.0.0
      directly attached to vlan2
N    172.16.100.8/30      [10] area: 0.0.0.0
      directly attached to vlan4
N    172.16.100.16/30     [20] area: 0.0.0.0
      via 172.16.100.5, vlan2
      via 172.16.100.9, vlan4
N IA 192.168.0.0/18    [20] area: 0.0.0.0
      via 172.16.100.9, vlan4
N    192.168.128.0/24    [10] area: 0.0.0.1
      directly attached to eth1
N    192.168.129.0/24    [10] area: 0.0.0.1
      directly attached to vlan3
N IA 192.168.192.0/18 [20] area: 0.0.0.0
      via 172.16.100.5, vlan2

===== OSPF router routing table =====
R    172.16.100.17        [10] area: 0.0.0.0, ABR
      via 172.16.100.9, vlan4
R    172.16.100.18        [10] area: 0.0.0.0, ABR
      via 172.16.100.5, vlan2

```

Notar que en la base de dades ja es reflecteix l'agregació de rutes de les subxarxes remotes (no de les locals):

```

OpenWrt# show ip ospf database

      OSPF Router with ID (172.16.100.10)

          Router Link States (Area 0.0.0.0)

Link ID        ADV Router      Age  Seq#           CkSum  Link count
172.16.100.10  172.16.100.10  1214 0x80000056     0xlcc3  2
172.16.100.17  172.16.100.17  1215 0x8000004d     0xac19  2
172.16.100.18  172.16.100.18  1210 0x80000002     0x8095  2

          Net Link States (Area 0.0.0.0)

Link ID        ADV Router      Age  Seq#           CkSum

```

```

172.16.100.6    172.16.100.10    1214 0x80000001 0x1795
172.16.100.10 172.16.100.10    1385 0x80000004 0xdec7
172.16.100.17 172.16.100.17    1215 0x80000001 0xe4ae

```

Summary Link States (Area 0.0.0.0)

Link ID	ADV Router	Age	Seq#	CkSum	Route
192.168.0.0	172.16.100.17	142	0x80000005	0x22d0	192.168.0.0/18
192.168.128.0	172.16.100.10	192	0x80000006	0xc4b3	192.168.128.0/18
192.168.192.0	172.16.100.18	1214	0x80000001	0xdb59	192.168.192.0/18

Router Link States (Area 0.0.0.1 [Stub])

Link ID	ADV Router	Age	Seq#	CkSum	Link count
172.16.100.10	172.16.100.10	144	0x80000007	0x936b	2

Summary Link States (Area 0.0.0.1 [Stub])

Link ID	ADV Router	Age	Seq#	CkSum	Route
0.0.0.0	172.16.100.10	1533	0x80000004	0xff30	0.0.0.0/0

Com fins ara, es generen un parell de captues sobre el rendiment de l'OSPF:

```

Mem: 10116K used, 4260K free, 0K shrd, 820K buff, 3436K cached
Load average: 0.00, 0.00, 0.00 (State: S=sleeping R=running, W=waiting)

```

```

  PID USER      STATUS  RSS  PPID  %CPU  %MEM  COMMAND
  4045 root        R       408   3939  1.9   2.8   top
  3615 quagga     S     1384     1  0.0   9.6   ospfd
  3603 quagga     S       852     1  0.0   5.9   zebra

```

A2.3.3. Router 2

La configuració de l'últim encaminador de l'escenari és la següent:

```

!
hostname OpenWRT@AP2
log stdout
!
interface eth0
!
interface eth1
 ip address 192.168.193.1/24
 ip ospf cost 10
!
interface lo
 ip address 127.0.0.1/8
!
interface vlan0
!
interface vlan1
 ip address 172.16.100.18/30
 ip ospf cost 10
!
interface vlan2
 ip address 172.16.100.5/30
 ip ospf cost 10
!
interface vlan3
 ip address 192.168.192.1/24
 ip ospf cost 10
!
interface vlan4
!
router ospf

```

```

ospf router-id 172.16.100.18
network 172.16.100.4/30 area 0.0.0.0
network 172.16.100.8/30 area 0.0.0.0
network 172.16.100.16/30 area 0.0.0.0
network 192.168.192.0/24 area 0.0.0.2
network 192.168.193.0/24 area 0.0.0.2
area 0.0.0.2 range 192.168.192.0/24 substitute 192.168.192.0/18
area 0.0.0.2 range 192.168.193.0/24 not-advertise
!
ip forwarding
ipv6 forwarding
!
line vty
!

```

Notar que, en la configuració de l'àrea 0.0.0.2 s'ha d'indicar implícitament els rangs que no es volen anunciar si s'ha sumartizat l'àrea. Com que l'identificador OSPF és el major de la xarxa, aquest equip pren el rol de Designated Router en els dos enllaços:

```
OpenWrt# show ip ospf neighbor
```

Neighbor ID	Pri	State	Dead Time	Address	Interface
RXmtL RqstL DBsmL					
172.16.100.17	1	Full/DR	00:00:36	172.16.100.17	
vlan1:172.16.100.18	0	0 0			
172.16.100.10	1	Full/DR	00:00:36	172.16.100.6	
vlan2:172.16.100.5	0	0 0			

Per acabar, les rutes que es generen són:

```

OpenWrt# show ip route
Codes: K - kernel route, C - connected, S - static, R - RIP, O - OSPF,
       I - ISIS, B - BGP, > - selected route, * - FIB route

C>* 127.0.0.0/8 is directly connected, lo
O 172.16.100.4/30 [110/10] is directly connected, vlan2, 00:14:23
C>* 172.16.100.4/30 is directly connected, vlan2
O>* 172.16.100.8/30 [110/20] via 172.16.100.17, vlan1, 00:14:13
       via 172.16.100.6, vlan2, 00:14:13
O 172.16.100.16/30 [110/10] is directly connected, vlan1, 00:14:23
C>* 172.16.100.16/30 is directly connected, vlan1
O>* 192.168.0.0/18 [110/20] via 172.16.100.17, vlan1, 00:14:13
O>* 192.168.128.0/18 [110/20] via 172.16.100.6, vlan2, 00:14:13
O 192.168.192.0/24 [110/10] is directly connected, vlan3, 00:14:23
C>* 192.168.192.0/24 is directly connected, vlan3
O 192.168.193.0/24 [110/10] is directly connected, eth1, 00:14:23
C>* 192.168.193.0/24 is directly connected, eth1

```

Com ja s'ha comentat, la mida de la base de dades de cada area es redueix donada l'agregació de rutes de l'OSPF.

```
OpenWrt# show ip ospf database
```

```

OSPF Router with ID (172.16.100.18)

Router Link States (Area 0.0.0.0)

Link ID          ADV Router      Age Seq#          CkSum Link count
172.16.100.10    172.16.100.10   943 0x80000056 0xlcc3 2
172.16.100.17    172.16.100.17   942 0x8000004d 0xac19 2
172.16.100.18    172.16.100.18   936 0x80000002 0x8095 2

Net Link States (Area 0.0.0.0)

```

```

Link ID      ADV Router    Age  Seq#      CkSum
172.16.100.6 172.16.100.10 942 0x80000001 0x1795
172.16.100.10 172.16.100.10 1114 0x80000004 0xdec7
172.16.100.17 172.16.100.17 942 0x80000001 0xe4ae

```

Summary Link States (Area 0.0.0.0)

```

Link ID      ADV Router    Age  Seq#      CkSum  Route
192.168.0.0 172.16.100.17 1721 0x80000004 0x24cf 192.168.0.0/18
192.168.128.0 172.16.100.10 1712 0x80000005 0xc6b2 192.168.128.0/18
192.168.192.0 172.16.100.18 941 0x80000001 0xdb59 192.168.192.0/18

```

Router Link States (Area 0.0.0.2)

```

Link ID      ADV Router    Age  Seq#      CkSum  Link count
172.16.100.18 172.16.100.18 899 0x80000003 0x78f7 2

```

Summary Link States (Area 0.0.0.2)

```

Link ID      ADV Router    Age  Seq#      CkSum  Route
172.16.100.4 172.16.100.18 941 0x80000001 0x02fb 172.16.100.4/30
172.16.100.8 172.16.100.18 931 0x80000001 0x3eb1 172.16.100.8/30
172.16.100.16 172.16.100.18 941 0x80000001 0x8968 172.16.100.16/30
192.168.0.0 172.16.100.18 931 0x80000001 0x8863 192.168.0.0/18
192.168.128.0 172.16.100.18 931 0x80000001 0x0368 192.168.128.0/18

```

I finalment, s'aprecia que el consum de recursos és el mateix que en la resta d'encaminadors.

```

Mem: 10504K used, 3872K free, 0K shrd, 608K buff, 4084K cached
Load average: 0.00, 0.00, 0.00 (State: S=sleeping R=running, W=waiting)

```

```

  PID USER      STATUS  RSS  PPID  %CPU  %MEM  COMMAND
 1390 root        S       648   437   0.7   4.5  dropbear
 1523 root        R       412  1391   0.5   2.8  top
 1490 quagga     S      1356    1   0.0   9.4  ospfd
 1487 quagga     S       856    1   0.0   5.9  zebra

```

Fins aquest moment no s'havia comentat que les rutes del quagga/bird s'exporten directament a la taula d'encaminament del Linux. Un exemple:

```

root@OpenWrt:/etc/quagga# netstat -r
Kernel IP routing table
Destination      Gateway          Genmask         Flags         MSS Window  irtt Iface
172.16.100.8     172.16.100.17  255.255.255.252 UG            0 0        0  vlan1
172.16.100.4     *                255.255.255.252 U              0 0        0  vlan2
172.16.100.16   *                255.255.255.252 U              0 0        0  vlan1
192.168.193.0   *                255.255.255.0  U             0 0        0  eth1
192.168.192.0   *                255.255.255.0  U             0 0        0  vlan3
192.168.128.0   172.16.100.6   255.255.192.0  UG            0 0        0  vlan2
192.168.0.0     172.16.100.17  255.255.192.0  UG            0 0        0  vlan1

```

A2.3.4. Test de tolerància a fallides

Partint de l'escenari estable es talla l'enllaç entre Router 3 i Router 4. Anteriorment, la taula d'encaminament (vista des de R3) és la que es mostra a continuació:

```

OpenWrt# show ip route ospf
Codes: K - kernel route, C - connected, S - static, R - RIP, O - OSPF,
       I - ISIS, B - BGP, > - selected route, * - FIB route

O   172.16.100.4/30 [110/10] is directly connected, vlan2, 02:10:10

```

```

O 172.16.100.8/30 [110/10] is directly connected, vlan4, 02:08:39
O>* 172.16.100.16/30 [110/20] via 172.16.100.5, vlan2, 00:27:16
    via 172.16.100.9, vlan4, 00:27:16
O>* 192.168.0.0/18 [110/20] via 172.16.100.9, vlan4, 02:00:08
O 192.168.128.0/24 [110/10] is directly connected, eth1, 02:10:14
O 192.168.129.0/24 [110/10] is directly connected, vlan3, 02:10:14
O>* 192.168.192.0/18 [110/20] via 172.16.100.5, vlan2, 00:27:16

```

Després de tallar l'enllaç el comptador de veïnatge no es renova i s'acaba esgotant:

```
OpenWrt# show ip ospf neighbor
```

Neighbor ID	Pri	State	Dead Time	Address	Interface
172.16.100.18	1	Full/Backup	00:00:33	172.16.100.5	172.16.100.5
vlan2:172.16.100.6	0	0 0			
172.16.100.17	1	Full/Backup	00:00:27	172.16.100.9	172.16.100.9
vlan4:172.16.100.10	0	0 0			

```
OpenWrt# show ip ospf neighbor
```

```
OpenWrt# show ip ospf neighbor
```

Neighbor ID	Pri	State	Dead Time	Address	Interface
172.16.100.18	1	Full/Backup	00:00:33	172.16.100.5	172.16.100.5
vlan2:172.16.100.6	0	0 0			
172.16.100.17	1	Full/Backup	00:00:07	172.16.100.9	172.16.100.9
vlan4:172.16.100.10	0	0 0			

```
OpenWrt# show ip ospf neighbor
```

Neighbor ID	Pri	State	Dead Time	Address	Interface
172.16.100.18	1	Full/Backup	00:00:40	172.16.100.5	172.16.100.5
vlan2:172.16.100.6	0	0 0			
172.16.100.17	1	Full/Backup	00:00:04	172.16.100.9	172.16.100.9
vlan4:172.16.100.10	0	0 0			

Amb el que s'ha de recalculer una nova ruta, el camí R3-R3-R4:

```
OpenWrt# show ip route ospf
```

Codes: K - kernel route, C - connected, S - static, R - RIP, O - OSPF,
I - ISIS, B - BGP, > - selected route, * - FIB route

```

O 172.16.100.4/30 [110/10] is directly connected, vlan2, 02:12:09
O 172.16.100.8/30 [110/10] is directly connected, vlan4, 02:10:38
O>* 172.16.100.16/30 [110/20] via 172.16.100.5, vlan2, 00:01:04
O>* 192.168.0.0/18 [110/30] via 172.16.100.5, vlan2, 00:00:47
O 192.168.128.0/24 [110/10] is directly connected, eth1, 02:12:13
O 192.168.129.0/24 [110/10] is directly connected, vlan3, 02:12:13
O>* 192.168.192.0/18 [110/20] via 172.16.100.5, vlan2, 00:29:15

```

Es comprova que el temps de convergència és menor a 1 minut, tant en la caiguda de l'enllaç com en la reconstrucció de la xarxa original un cop l'enllaç torna a estar disponible.

A2.4. Escenari 4

El quart escenari és una extensió de l'escenari tercer. S'ha configurat en enllaç redundat entre dos dels encaminadors. En aquest escenari es volen realitzar dos tests diferents:

- Tornar a veure el correcte funcionament de la tolerància a fallides.
- Veure el funcionament del balanceig de càrrega.

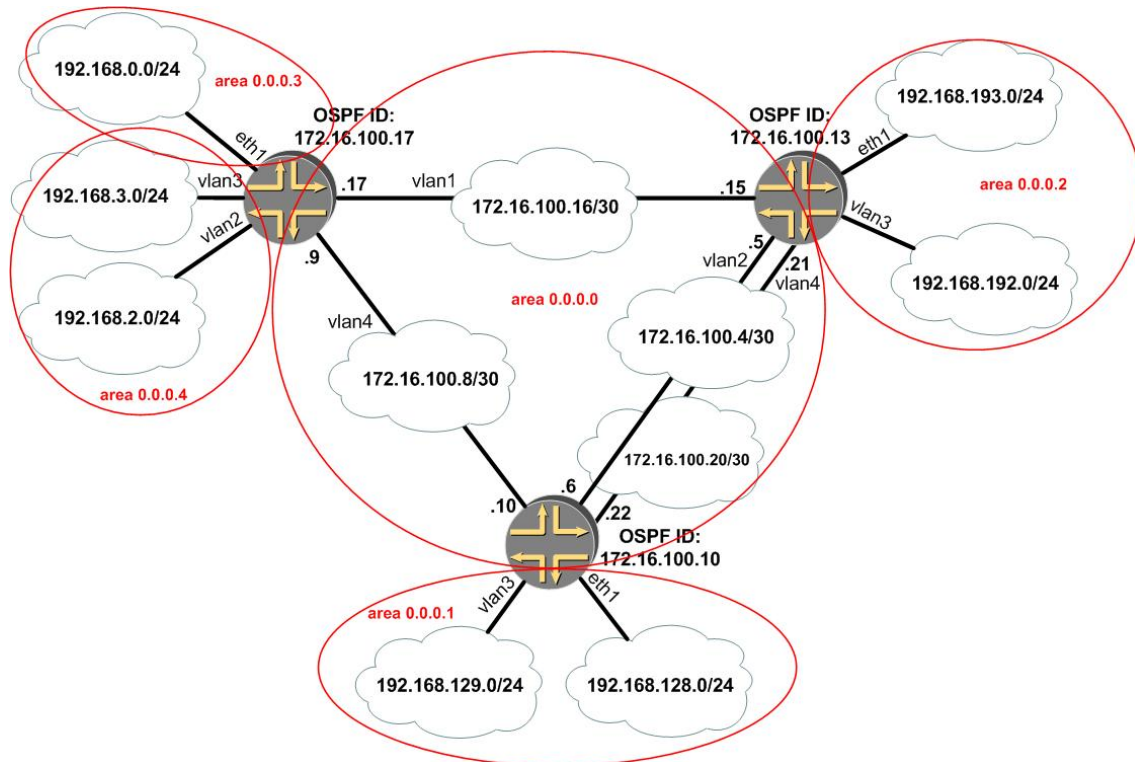


Fig. A2.4 Quart escenari OSPF.

A2.4.1. Router 4

La configuració del router 4, amb OSPF ID 172.16.100.17 és la següent:

```
!
hostname OpenWRT@AP4
log stdout
!
password ap4
enable password ap4
!
interface eth0
!
interface eth1
 ip address 192.168.0.1/24
 ip ospf cost 10
!
interface lo
 ip address 127.0.0.1/8
!
interface vln0
```



```

ip address 192.168.1.1/24
no ip ospf
!
interface vlan1
ip address 172.16.100.17/30
ip ospf cost 10
!
interface vlan2
ip address 192.168.2.1/24
ip ospf cost 10
!
interface vlan3
ip address 192.168.3.1/24
ip ospf cost 10
!
interface vlan4
ip address 172.16.100.9/30
ip ospf cost 10
!
router ospf
ospf router-id 172.16.100.17
network 172.16.100.0/24 area 0.0.0.0
network 192.168.0.0/24 area 0.0.0.3
network 192.168.2.0/24 area 0.0.0.4
network 192.168.3.0/24 area 0.0.0.4
area 0.0.0.3 range 192.168.0.0/24 substitute 192.168.0.0/18
area 0.0.0.4 range 192.168.2.0/24 not-advertise
area 0.0.0.4 range 192.168.3.0/24 not-advertise
!
ip forwarding
ipv6 forwarding
!
line vty
!

```

I es reflecteix amb la creació de les següents rutes:

```

OpenWrt# show ip ospf route
===== OSPF network routing table =====
N    172.16.100.4/30      [20] area: 0.0.0.0
      via 172.16.100.18, vlan1
      via 172.16.100.10, vlan4
N    172.16.100.8/30      [10] area: 0.0.0.0
      directly attached to vlan4
N    172.16.100.16/30     [10] area: 0.0.0.0
      directly attached to vlan1
N    172.16.100.20/30    [25] area: 0.0.0.0
      via 172.16.100.18, vlan1
      via 172.16.100.10, vlan4
N    192.168.0.0/24     [10] area: 0.0.0.3
      directly attached to eth1
N    192.168.2.0/24     [10] area: 0.0.0.4
      directly attached to vlan2
N    192.168.3.0/24     [10] area: 0.0.0.4
      directly attached to vlan3
N IA 192.168.128.0/18  [20] area: 0.0.0.0
      via 172.16.100.10, vlan4
N IA 192.168.192.0/18 [20] area: 0.0.0.0
      via 172.16.100.18, vlan1

===== OSPF router routing table =====
R    172.16.100.10      [10] area: 0.0.0.0, ABR
      via 172.16.100.10, vlan4
R    172.16.100.18      [10] area: 0.0.0.0, ABR
      via 172.16.100.18, vlan1

```

El consum de recursos és el mateix que en els altres escenaris, tot i que en aquest s'està encaminant tot el rang 192.168.0.0/16 a més de les adreces de troncal. L'aplicació vtysh, per accedir a la consola del quagga consumeix molts recursos. Es recomana tancar-la al abandonar l'administració de l'equip.

```
Mem: 11984K used, 2392K free, 0K shrd, 820K buff, 3372K cached
Load average: 0.05, 0.07, 0.02 (State: S=sleeping R=running, W=waiting)
```

PID	USER	STATUS	RSS	PPID	%CPU	%MEM	COMMAND
2483	root	R	408	2423	2.8	2.8	top
2177	root	S	3040	2159	0.0	21.1	vtys
2459	quagga	S	1356	1	0.0	9.4	ospfd
2406	quagga	S	860	1	0.0	5.9	zebra

A2.4.2. Router 3

La configuració del router 3 és la següent. Notar les diferents mètriques aplicades a les interfícies. Aquesta configuració correspon a la de la primera part de l'escenari, on l'enllaç redundat realitza funcions de backup.

```
Current configuration:
!
hostname OpenWRT@AP3
log stdout
!
debug ospf event
debug ospf packet all
!
interface eth0
!
interface eth1
 ip address 192.168.128.1/24
 ip ospf cost 10
!
interface lo
 ip address 127.0.0.1/8
!
interface vlan0
 ip address 192.168.1.1/24
!
interface vlan1
 ip address 172.16.100.22/30
 ip ospf cost 15
!
interface vlan2
 ip address 172.16.100.6/30
 ip ospf cost 10
!
interface vlan3
 ip address 192.168.129.1/24
 ip ospf cost 10
!
interface vlan4
 ip address 172.16.100.10/30
 ip ospf cost 10
!
router ospf
 ospf router-id 172.16.100.10
 network 172.16.100.0/24 area 0.0.0.0
 network 192.168.128.0/24 area 0.0.0.1
 network 192.168.129.0/24 area 0.0.0.1
```

```

area 0.0.0.1 stub no-summary
area 0.0.0.1 range 192.168.128.0/24 substitute 192.168.128.0/18
area 0.0.0.1 range 192.168.129.0/24 not-advertise
!
ip forwarding
ipv6 forwarding
!
line vty
!

```

El procés de veinatge es realitza correctament. Notar que ara l'equip té 3 veïns en comptes de 2:

```
OpenWrt# show ip ospf neighbor
```

Neighbor	ID	Pri	State	Dead Time	Address	Interface
172.16.100.18		1	Full/Backup	00:00:30	172.16.100.21	vlan1:172.16.100.22
0	0	0				
172.16.100.18		1	Full/Backup	00:00:30	172.16.100.5	vlan2:172.16.100.6
0	0	0				
172.16.100.17		1	Full/Backup	00:00:39	172.16.100.9	vlan4:172.16.100.10
0	0	0				

Aparentment les rutes a les xarxes d'usuari no es veuen alterades, però justament es pot observar que la taula d'encaminament ja mostra el camí alternatiu:

```
OpenWrt# show ip route
```

```
Codes: K - kernel route, C - connected, S - static, R - RIP, O - OSPF,
I - ISIS, B - BGP, > - selected route, * - FIB route
```

```

C>* 127.0.0.0/8 is directly connected, lo
O 172.16.100.4/30 [110/10] is directly connected, vlan2, 00:29:09
C>* 172.16.100.4/30 is directly connected, vlan2
O 172.16.100.8/30 [110/10] is directly connected, vlan4, 00:05:34
C>* 172.16.100.8/30 is directly connected, vlan4
O>* 172.16.100.16/30 [110/20] via 172.16.100.5, vlan2, 00:05:34
via 172.16.100.9, vlan4, 00:05:34
O 172.16.100.20/30 [110/15] is directly connected, vlan1, 00:29:04
C>* 172.16.100.20/30 is directly connected, vlan1
O>* 192.168.0.0/18 [110/20] via 172.16.100.9, vlan4, 00:05:34
C>* 192.168.1.0/24 is directly connected, vlan0
O 192.168.128.0/24 [110/10] is directly connected, eth1, 00:32:20
C>* 192.168.128.0/24 is directly connected, eth1
O 192.168.129.0/24 [110/10] is directly connected, vlan3, 00:32:20
C>* 192.168.129.0/24 is directly connected, vlan3
O>* 192.168.192.0/18 [110/20] via 172.16.100.5, vlan2, 00:29:04

```

```
OpenWrt# show ip ospf route
```

```

===== OSPF network routing table =====
N 172.16.100.4/30 [10] area: 0.0.0.0
directly attached to vlan2
N 172.16.100.8/30 [10] area: 0.0.0.0
directly attached to vlan4
N 172.16.100.16/30 [20] area: 0.0.0.0
via 172.16.100.5, vlan2
via 172.16.100.9, vlan4
N 172.16.100.20/30 [15] area: 0.0.0.0
directly attached to vlan1
N IA 192.168.0.0/18 [20] area: 0.0.0.0
via 172.16.100.9, vlan4
N 192.168.128.0/24 [10] area: 0.0.0.1
directly attached to eth1
N 192.168.129.0/24 [10] area: 0.0.0.1
directly attached to vlan3

```

```

N IA 192.168.192.0/18      [20] area: 0.0.0.0
                           via 172.16.100.5, vlan2

===== OSPF router routing table =====
R   172.16.100.17         [10] area: 0.0.0.0, ABR
                           via 172.16.100.9, vlan4
R   172.16.100.18         [10] area: 0.0.0.0, ABR
                           via 172.16.100.5, vlan2

```

El consum de recursos ha augmentat un 0.4% en el cas de la memòria:

```

Mem: 11044K used, 3332K free, 0K shrd, 828K buff, 4532K cached
Load average: 0.00, 0.00, 0.00    (State: S=sleeping R=running, W=waiting)

```

PID	USER	STATUS	RSS	PPID	%CPU	%MEM	COMMAND
1182	root	R	408	1162	1.9	2.8	top
1161	root	S	648	505	0.9	4.5	dropbear
1088	quagga	S	1396	1	0.0	9.7	ospfd
1076	quagga	S	844	1	0.0	5.8	zebra

A2.4.3. Router 2

La configuració del tercer equip, que també reflecteix les diferents mètriques aplicades a les interfícies:

```

Current configuration:
!
hostname OpenWRT@AP2
log stdout
!
interface eth0
!
interface eth1
 ip address 192.168.193.1/24
 ip ospf cost 10
!
interface lo
 ip address 127.0.0.1/8
!
interface vlan0
!
interface vlan1
 ip address 172.16.100.18/30
 ip ospf cost 10
!
interface vlan2
 ip address 172.16.100.5/30
 ip ospf cost 10
!
interface vlan3
 ip address 192.168.192.1/24
 ip ospf cost 10
!
interface vlan4
 ip address 172.16.100.21/30
 ip ospf cost 15
!
router ospf
 ospf router-id 172.16.100.18
 network 172.16.100.0/24 area 0.0.0.0
 network 192.168.192.0/24 area 0.0.0.2
 network 192.168.193.0/24 area 0.0.0.2
 area 0.0.0.2 range 192.168.192.0/24 substitute 192.168.192.0/18
 area 0.0.0.2 range 192.168.193.0/24 not-advertise

```

```
!
ip forwarding
ipv6 forwarding
!
line vty
!
```

El nombre de veïns continua essent 3:

```
OpenWrt# show ip ospf neighbor
```

Neighbor ID	Pri	State	Dead Time	Address	Interface
172.16.100.17	1	Full/Backup	00:00:34		172.16.100.17
vlan1:172.16.100.18	0	0 0			
172.16.100.10	1	Full/DR	00:00:35		172.16.100.6
vlan2:172.16.100.5	0	0 0			
172.16.100.10	1	Full/DR	00:00:35		172.16.100.22
vlan4:172.16.100.21	0	0 0			

I els recursos es mantenen amb els mateixos valors que fins al moment:

```
Mem: 10132K used, 4244K free, 0K shrd, 820K buff, 3512K cached
Load average: 1.09, 0.99, 0.79 (State: S=sleeping R=running, W=waiting)
```

PID	USER	STATUS	RSS	PPID	%CPU	%MEM	COMMAND
1158	root	R	396	1	97.6	2.7	vi
1274	root	R	408	1255	2.9	2.8	top
1173	quagga	S	1372	1	0.0	9.5	ospfd
1170	quagga	S	860	1	0.0	5.9	zebra

A2.4.4. Test de tolerància a fallides

Els veïns OSPF mostren que realment si existeix un camí redundant tot i no aparèixer la doble ruta cap a 192.168.128.0/24, per exemple. Com es pot veure, el camí al que s'ha configurat un enllaç de backup són les vlans2 i 4, sense poder distingir fins al moment quin és el camí preferent:

```
OpenWrt# show ip ospf neighbor
```

Neighbor ID	Pri	State	Dead Time	Address	Interface
172.16.100.17	1	Full/DR	00:00:37		172.16.100.17
vlan1:172.16.100.18	0	0 0			
172.16.100.10	1	Full/DR	00:00:37		172.16.100.6
vlan2:172.16.100.5	0	0 0			
172.16.100.10	1	Full/DR	00:00:37		172.16.100.22
vlan4:172.16.100.21	0	0 0			

La ruta habitual (metrica 20) és la que segueix la vlan2 i per tant, la vlan4 resta inactiva.

```
OpenWrt# show ip route 192.168.128.1
Routing entry for 192.168.128.0/18
  Known via "ospf", distance 110, metric 20, best
  Last update 00:00:36 ago
  * 172.16.100.6, via vlan2
```

Després de tallar l'enllaç de la vlan2 el comptador de veïnatge es va esgotant, sense que la vlan4 es vegi alterada.

Neighbor ID	Pri	State	Dead Time	Address	Interface
RXmtL RqstL DBsmL					
172.16.100.17	1	Full/DR		00:00:38	172.16.100.17
vlan1:172.16.100.18	0	0 0			
172.16.100.10	1	Full/DR		00:00:18	172.16.100.6
vlan2:172.16.100.5	0	0 0			
172.16.100.10	1	Full/DR		00:00:38	172.16.100.22
vlan4:172.16.100.21	0	0 0			

Quan s'ha esgotat el comptador, el veí només és accessible per la vlan4:

Neighbor ID	Pri	State	Dead Time	Address	Interface
RXmtL RqstL DBsmL					
172.16.100.17	1	Full/DR		00:00:39	172.16.100.17
vlan1:172.16.100.18	0	0 0			
172.16.100.10	1	Full/DR		00:00:39	172.16.100.22
vlan4:172.16.100.21	0	0 0			

El que produeix que es segueixi el camí de backup (mètrica 25):

```
Ara s'utilitza l'enllaç redundat
OpenWrt# show ip route 192.168.128.1
Routing entry for 192.168.128.0/18
  Known via "ospf", distance 110, metric 25, best
  Last update 00:04:06 ago
  * 172.16.100.22, via vlan4
```

Si es restableix l'enllaç, i donat que es tracta d'un backup, el tràfic torna a circular per la vlan2:

```
OpenWrt# show ip route 192.168.128.1
Routing entry for 192.168.128.0/18
  Known via "ospf", distance 110, metric 20, best
  Last update 00:00:00 ago
  * 172.16.100.6, via vlan2
```

A2.4.5. Test de balanceig de càrrega

Si es modifiquen les mètriques amb valor 15 de tota la configuració i se'ls assigna un pes 10 els dos enllaços redundants passen a tenir la mateixa preferència:

Anteriorment només apareix una ruta seleccionada

```
OpenWrt# show ip route 192.168.128.1
Routing entry for 192.168.128.0/18
  Known via "ospf", distance 110, metric 20, best
  Last update 00:12:45 ago
  * 172.16.100.6, via vlan2
```

Mentre que ara apareixen els dos camins com actius:

```
OpenWrt# show ip route 192.168.128.1
Routing entry for 192.168.128.0/18
  Known via "ospf", distance 110, metric 20, best
  Last update 00:03:28 ago
  * 172.16.100.6, via vlan2
  * 172.16.100.22, via vlan4
```

Després de generar tràfic es comprova que per ambdues interfícies ha passat aproximadament la mateixa quantitat d'informació.

A2.5. Escenari 5

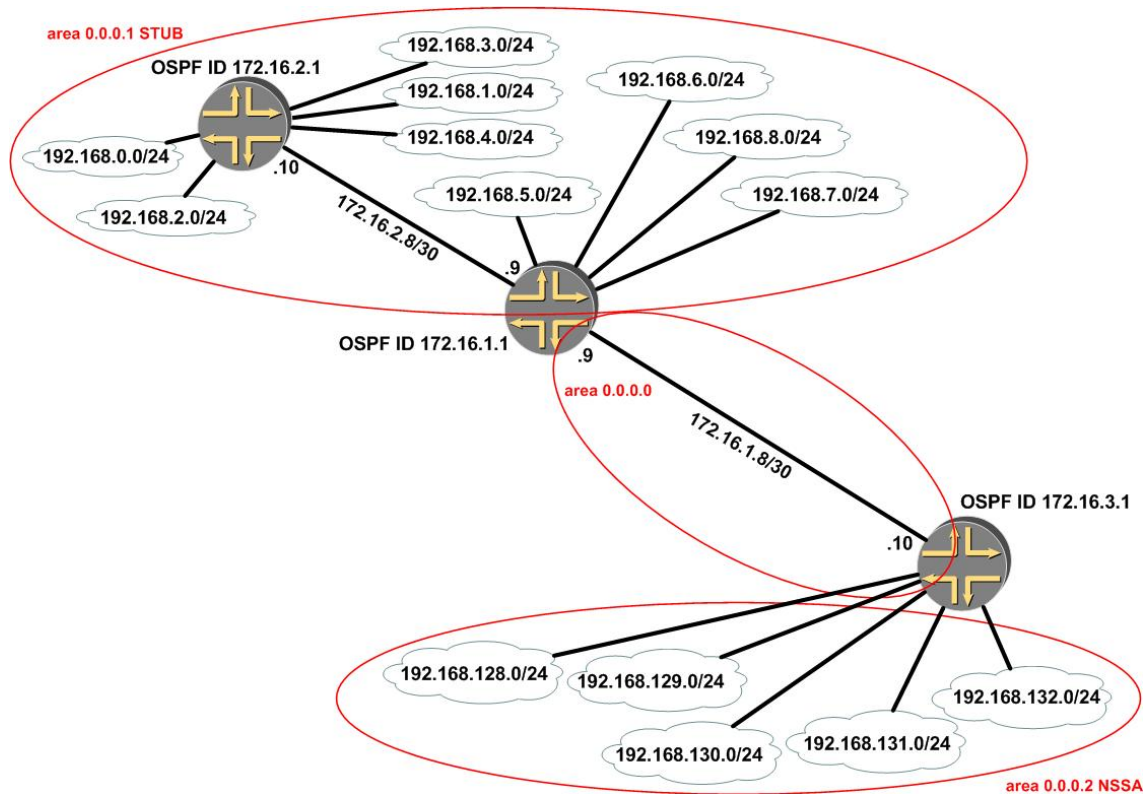


Fig. A2.5 Cinquè escenari d'OSPF

A2.5.1. Router 4

La configuració del router 4, amb OSPF ID 172.16.3.1 és:

```
hostname OpenWRT@AP4
log stdout
!
password ap4
enable password ap4
!
interface eth0
!
interface eth1
ip address 192.168.132.1/24
!
interface lo
ip address 127.0.0.1/8
!
interface vlan0
ip address 192.168.128.1/24
!
interface vlan1
ip address 172.16.1.10/30
ip ospf cost 10
!
```

```

interface vlan2
 ip address 192.168.129.1/24
 ip ospf cost 10
!
interface vlan3
 ip address 192.168.130.1/24
 ip ospf cost 10
!
interface vlan4
 ip address 192.168.131.1/24
 ip ospf cost 10
! router ospf
 ospf router-id 172.16.3.1
 network 172.16.1.8/30 area 0.0.0.0
 network 192.168.128.0/24 area 0.0.0.2
 network 192.168.129.0/24 area 0.0.0.2
 network 192.168.130.0/24 area 0.0.0.2
 network 192.168.131.0/24 area 0.0.0.2
 network 192.168.132.0/24 area 0.0.0.2
 area 0.0.0.2 nssa translate-candidate
 area 0.0.0.2 range 192.168.128.0/24 substitute 192.168.128.0/21
 area 0.0.0.2 range 192.168.129.0/24 not-advertise
 area 0.0.0.2 range 192.168.130.0/24 not-advertise
 area 0.0.0.2 range 192.168.131.0/24 not-advertise
 area 0.0.0.2 range 192.168.132.0/24 not-advertise
 default-information originate always
!
 ip forwarding
 ipv6 forwarding
!
 line vty
!

```

Amb el qual, i com ja s'ha comentat, l'àrea 0.0.0.2 es configura com NSSA i l'únic encaminador (l'ABR) pren el rol de traductor:

```

Area ID: 0.0.0.2 (NSSA)
  Shortcutting mode: Default, S-bit consensus: ok
  Number of interfaces in this area: Total: 5, Active: 5
  It is an NSSA configuration.
  Elected NSSA/ABR performs type-7/type-5 LSA translation.
  We are an ABR and the NSSA Elected Translator.

```

Aquest cop la base de dades de les àrees connectades a aquest equip és la següent:

```
OpenWrt# show ip ospf database
```

```

      OSPF Router with ID (172.16.3.1)

          Router Link States (Area 0.0.0.0)

Link ID        ADV Router    Age  Seq#           CkSum  Link count
172.16.1.1    172.16.1.1    950  0x80000052    0xf4f6  1
172.16.3.1    172.16.3.1    867  0x8000004e    0xecfb  1

          Net Link States (Area 0.0.0.0)

Link ID        ADV Router    Age  Seq#           CkSum
172.16.1.10    172.16.3.1    954  0x80000003    0x76dc

          Summary Link States (Area 0.0.0.0)

Link ID        ADV Router    Age  Seq#           CkSum  Route
172.16.2.8     172.16.1.1    1315 0x8000003b    0xbd8  172.16.2.8/30
192.168.0.0    172.16.1.1    103  0x80000004    0x919b  192.168.0.0/20
192.168.128.0  172.16.3.1    191  0x8000003b    0x5322  192.168.128.0/21

```



```

Router Link States (Area 0.0.0.2 [NSSA])

Link ID          ADV Router      Age Seq#          CkSum Link count
172.16.3.1      172.16.3.1      867 0x8000000a 0x1899 5

Summary Link States (Area 0.0.0.2 [NSSA])

Link ID          ADV Router      Age Seq#          CkSum Route
0.0.0.0         172.16.3.1     1473 0x80000003 0xe1b9 0.0.0.0/0
172.16.1.8     172.16.3.1      31 0x80000004 0x4787 172.16.1.8/30
172.16.2.8     172.16.3.1     542 0x80000003 0xa222 172.16.2.8/30
192.168.0.0    172.16.3.1     1543 0x80000003 0x081c 192.168.0.0/20

NSSA-external Link States (Area 0.0.0.2 [NSSA])

Link ID          ADV Router      Age Seq#          CkSum Route
0.0.0.0         172.16.3.1     292 0x80000003 0x1fed E2 0.0.0.0/0 [0x0]

AS External Link States

Link ID          ADV Router      Age Seq#          CkSum Route
0.0.0.0         172.16.3.1     292 0x80000003 0x3bd3 E2 0.0.0.0/0 [0x0]

```

I es reflecteix amb la creació de les següents rutes. Notar que s'ha sumariat correctament l'area 0.0.0.1

```

OpenWrt# show ip ospf route
===== OSPF network routing table =====
N   172.16.1.8/30          [10] area: 0.0.0.0
    directly attached to vlan1
N IA 172.16.2.8/30       [20] area: 0.0.0.0
    via 172.16.1.9, vlan1
N IA 192.168.0.0/20     [30] area: 0.0.0.0
    via 172.16.1.9, vlan1
N   192.168.128.0/24     [10] area: 0.0.0.2
    directly attached to vlan0
N   192.168.129.0/24     [10] area: 0.0.0.2
    directly attached to vlan2
N   192.168.130.0/24     [10] area: 0.0.0.2
    directly attached to vlan3
N   192.168.131.0/24     [10] area: 0.0.0.2
    directly attached to vlan4
N   192.168.132.0/24     [10] area: 0.0.0.2
    directly attached to eth1
===== OSPF router routing table =====
R   172.16.1.1           [10] area: 0.0.0.0, ABR
    via 172.16.1.9, vlan1

```

El consum de recursos és el mateix que en els altres escenaris:

```

Mem: 10404K used, 3972K free, 0K shrd, 820K buff, 3900K cached
Load average: 0.00, 0.00, 0.00 (State: S=sleeping R=running, W=waiting)

```

```

  PID USER      STATUS  RSS  PPID  %CPU  %MEM  COMMAND
 3657 root        R       408  3629   1.9   2.8  top
 3493 quagga     S      1376    1   0.0   9.5  ospfd
 3481 quagga     S       876    1   0.0   6.0  zebra

```

A2.5.2. Router 3

La configuració del router 3, amb OSPF ID 172.16.1.1 :

```
Current configuration:
!
hostname OpenWRT@AP3
log stdout
!
debug ospf event
debug ospf packet all
!
interface eth0
!
interface eth1
 ip address 192.168.8.1/24
 ip ospf cost 10
!
interface lo
 ip address 127.0.0.1/8
!
interface vlan0
 ip address 192.168.5.1/24
!
interface vlan1
 ip address 172.16.1.9/30
 ip ospf cost 10
!
interface vlan2
 ip address 172.16.2.9/30
 ip ospf cost 10
!
interface vlan3
 ip address 192.168.6.1/24
 ip ospf cost 10
!
interface vlan4
 ip address 192.168.7.1/24
 ip ospf cost 10
!
router ospf
 ospf router-id 172.16.1.1
 network 172.16.1.8/30 area 0.0.0.0
 network 172.16.2.8/30 area 0.0.0.1
 network 192.168.5.0/24 area 0.0.0.1
 network 192.168.6.0/24 area 0.0.0.1
 network 192.168.7.0/24 area 0.0.0.1
 network 192.168.8.0/24 area 0.0.0.1
 area 0.0.0.1 stub no-summary
 area 0.0.0.1 range 192.168.0.0/24 substitute 192.168.0.0/20
 area 0.0.0.1 range 192.168.1.0/24 not-advertise
 area 0.0.0.1 range 192.168.2.0/24 not-advertise
 area 0.0.0.1 range 192.168.3.0/24 not-advertise
 area 0.0.0.1 range 192.168.4.0/24 not-advertise
 area 0.0.0.1 range 192.168.5.0/24 not-advertise
 area 0.0.0.1 range 192.168.6.0/24 not-advertise
 area 0.0.0.1 range 192.168.7.0/24 not-advertise
 area 0.0.0.1 range 192.168.8.0/24 not-advertise
!
ip forwarding
ipv6 forwarding
!
line vty
!
```

Les rutes queden de la següent manera. Notar que s'ha sumartizat correctament l'àrea 0.0.0.2:

```
OpenWrt# show ip route ospf
Codes: K - kernel route, C - connected, S - static, R - RIP, O - OSPF,
```

```

I - ISIS, B - BGP, > - selected route, * - FIB route

O 172.16.1.8/30 [110/10] is directly connected, vlan1, 01:19:34
O 172.16.2.8/30 [110/10] is directly connected, vlan2, 01:19:39
O>* 192.168.0.0/24 [110/20] via 172.16.2.10, vlan2, 01:19:35
O>* 192.168.1.0/24 [110/20] via 172.16.2.10, vlan2, 01:19:35
O>* 192.168.2.0/24 [110/20] via 172.16.2.10, vlan2, 01:19:35
O>* 192.168.3.0/24 [110/20] via 172.16.2.10, vlan2, 01:19:35
O>* 192.168.4.0/24 [110/20] via 172.16.2.10, vlan2, 01:19:35
O 192.168.5.0/24 [110/10] is directly connected, vlan0, 01:19:39
O 192.168.6.0/24 [110/10] is directly connected, vlan3, 01:19:39
O 192.168.7.0/24 [110/10] is directly connected, vlan4, 01:19:39
O 192.168.8.0/24 [110/10] is directly connected, eth1, 01:19:39
O>* 192.168.128.0/21 [110/20] via 172.16.1.10, vlan1, 00:01:39

```

El consum de recursos ha augmentat un 0.4% en el cas de la memòria:

```

Mem: 9776K used, 4600K free, 0K shrd, 736K buff, 3224K cached
Load average: 0.02, 0.02, 0.00 (State: S=sleeping R=running, W=waiting)

```

PID	USER	STATUS	RSS	PPID	%CPU	%MEM	COMMAND
4619	root	R	408	4478	2.8	2.8	top
4369	quagga	S	1392	1	0.0	9.6	ospfd
4357	quagga	S	868	1	0.0	6.0	zebra

A2.5.3. Router 2

La configuració del tercer equip, el router 2 amb OSPF ID 172.16.2.1 és:

```

Current configuration:
!
hostname OpenWRT@AP2
log stdout
!
interface eth0
!
interface eth1
ip address 192.168.4.1/24
ip ospf cost 10
!
interface lo
ip address 127.0.0.1/8
!
interface vlan0
ip address 192.168.1.1/24
ip ospf cost 10
!
interface vlan1
ip address 192.168.0.1/24
ip ospf cost 10
!
interface vlan2
ip address 172.16.2.10/30
ip ospf cost 10
!
interface vlan3
ip address 192.168.2.1/24
ip ospf cost 10
!
interface vlan4
ip address 192.168.3.1/24
ip ospf cost 10
!
router ospf
ospf router-id 172.16.2.1
network 172.16.2.8/30 area 0.0.0.1

```

```

network 192.168.0.0/24 area 0.0.0.1
network 192.168.1.0/24 area 0.0.0.1
network 192.168.2.0/24 area 0.0.0.1
network 192.168.3.0/24 area 0.0.0.1
network 192.168.4.0/24 area 0.0.0.1
area 0.0.0.1 stub no-summary
!
ip forwarding
ipv6 forwarding
!
line vty
!

```

l la corresponent base de dades generada i extremadament petita. Tot l'encaminament es basa en la ruta per defecte, injectada a l'àrea stub:

```
OpenWrt# show ip ospf database
```

```

OSPF Router with ID (172.16.2.1)

Router Link States (Area 0.0.0.1 [Stub])

Link ID          ADV Router      Age Seq#          CkSum Link count
172.16.1.1      172.16.1.1     1458 0x8000004e 0xc7ff 5
172.16.2.1      172.16.2.1     1496 0x8000004e 0x272b 6

Net Link States (Area 0.0.0.1 [Stub])

Link ID          ADV Router      Age Seq#          CkSum
172.16.2.10     172.16.2.1     1500 0x80000003 0x83d2

Summary Link States (Area 0.0.0.1 [Stub])

Link ID          ADV Router      Age Seq#          CkSum Route
0.0.0.0          172.16.1.1     539 0x8000003c 0x7de6 0.0.0.0/0

```

Les taules d'encaminament també son petites després de donar-hi un cop d'ull. La longitud es relativament gran ja que l'àrea 0.0.0.1 ho és, però la quantitat de rutes externes es redueix a una, la ruta per defecte. Notar que ni tant sols apareix la ruta cap a l'àrea 0.0.0.2, encara que segueix essent accessible sense cap problema:

```
OpenWrt# show ip route ospf
```

```
Codes: K - kernel route, C - connected, S - static, R - RIP, O - OSPF,
I - ISIS, B - BGP, > - selected route, * - FIB route
```

```

O>* 0.0.0.0/0 [110/11] via 172.16.2.9, vlan2, 01:25:36
O 172.16.2.8/30 [110/10] is directly connected, vlan2, 1d05h50m
O 192.168.1.0/24 [110/10] is directly connected, vlan0, 1d05h50m
O 192.168.2.0/24 [110/10] is directly connected, vlan3, 1d05h50m
O 192.168.3.0/24 [110/10] is directly connected, vlan4, 1d05h50m
O 192.168.4.0/24 [110/10] is directly connected, eth1, 1d05h50m
O>* 192.168.5.0/24 [110/20] via 172.16.2.9, vlan2, 01:25:36
O>* 192.168.6.0/24 [110/20] via 172.16.2.9, vlan2, 01:25:36
O>* 192.168.7.0/24 [110/20] via 172.16.2.9, vlan2, 01:25:36
O>* 192.168.8.0/24 [110/20] via 172.16.2.9, vlan2, 01:25:36

```


ANNEX 3: CONFIGURACIONS BGP

A3.1. Introducció

L'objectiu d'aquest annex és continuar amb la tasca de l'anterior, però aplicat a BGP. Per tant, s'aporten les configuracions i algun comentari sobre els punts estudiats en cada escenari.

A3.2. Escenari 1

Com a recordatori, en l'escenari primer es vol estudiar:

- Integració de OSPF i BGP. La part OSPF ha d'esser robusta.
- Comportament de BGP com a protocol exterior.
- Atributs del BGP.
- Possible control (filtrat) de les xarxes anunciades.

I per tant, es realitza un escenari amb 2 AS diferents i configuracions senzilles de BGP i robustes d'OSPF, sobretot en el cas de les àrees de l'AS100.

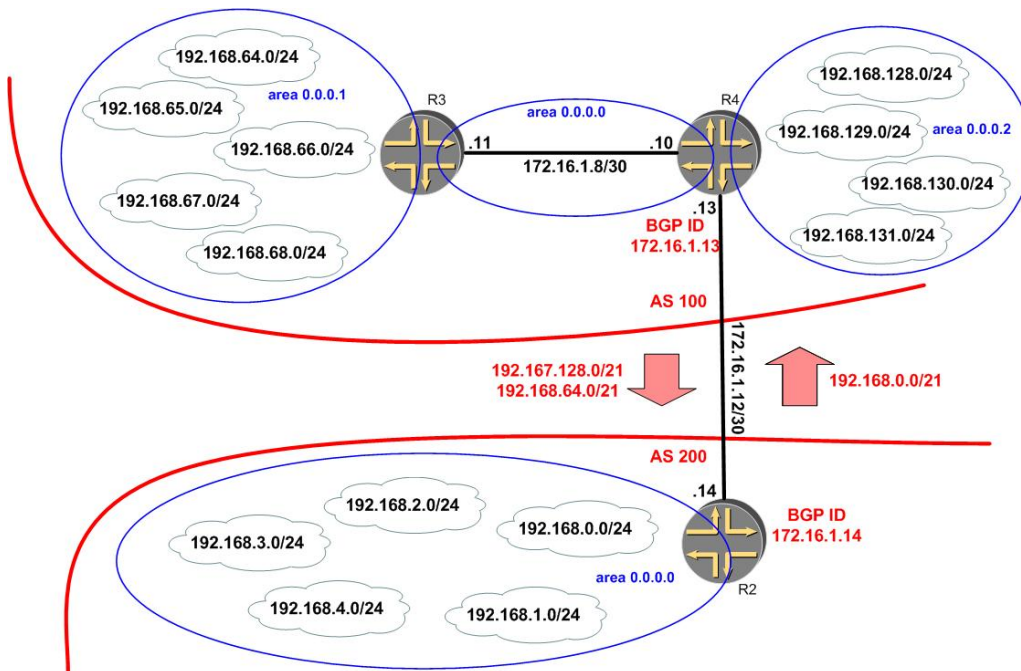


Fig. A3.1 Primer escenari de BGP

A3.2.1. Router 4

La configuració de l'equip que injecta les rutes externes del BGP dins l'IGP (OSPF) és la següent. Notar la sumarització de les rutes OSPF i com es controla la seva publicació dins el BGP.

```

Current configuration:
!
hostname OpenWRT@AP4
log stdout
!
password ap4
enable password ap4
!
interface eth0
!
interface eth1
 ip address 192.168.131.1/24
 ip ospf cost 10
!
interface lo
 ip address 127.0.0.1/8
!
interface vlan0
 ip address 192.168.128.1/24
 ip ospf cost 10
!
interface vlan1
 ip address 172.16.1.10/30
 ip ospf cost 10
!
interface vlan2
 ip address 192.168.129.1/24
 ip ospf cost 10
!
interface vlan3
 ip address 192.168.130.1/24
 ip ospf cost 10
!
interface vlan4
 ip address 172.16.1.13/30
!
router bgp 100
 bgp router-id 172.16.1.13
 network 172.16.1.12/30
 network 192.168.128.0/21
 redistribute ospf
 neighbor 172.16.1.14 remote-as 200
!
router ospf
 ospf router-id 172.16.3.1
 redistribute bgp
 passive-interface eth1
 network 172.16.1.8/30 area 0.0.0.0
 network 192.168.128.0/24 area 0.0.0.2
 network 192.168.129.0/24 area 0.0.0.2
 network 192.168.130.0/24 area 0.0.0.2
 network 192.168.131.0/24 area 0.0.0.2
 area 0.0.0.2 range 192.168.128.0/24 substitute 192.168.128.0/21
 area 0.0.0.2 range 192.168.129.0/24 not-advertise
 area 0.0.0.2 range 192.168.130.0/24 not-advertise
 area 0.0.0.2 range 192.168.131.0/24 not-advertise
 default-information originate always metric 30
!

```

Mitjançant les característiques i rol de l'OSPF ja es pot notar que l'encaminador realitza funcions d'ABR i també l'ASBR, injectant concretament 2 LSAs externs. Un d'aquest dos, correspon a la xarxa anunciada pel BGP corresponent a l'AS veí.

```
OpenWrt# show ip ospf
OSPF Routing Process, Router ID: 172.16.3.1
Supports only single TOS (TOS0) routes
This implementation conforms to RFC2328
RFC1583Compatibility flag is disabled
SPF schedule delay 1 secs, Hold time between two SPF's 1 secs
Refresh timer 10 secs
This router is an ABR, ABR type is: Standard (RFC2328)
This router is an ASBR (injecting external routing information)
Number of external LSA 2. Checksum Sum 0x00018d82
Number of areas attached to this router: 2
...
```

Com a curiositat, veure la base de dades OSPF, on encara no es veu la ruta injectada pel BGP, però si el sumari del BGP, que permet reduir el nombre de rutes dins la xarxa interna:

```
===== OSPF network routing table =====
N   172.16.1.8/30      [10] area: 0.0.0.0
                        directly attached to vlan1
N IA 192.168.64.0/21  [20] area: 0.0.0.0
                        via 172.16.1.9, vlan1
N   192.168.128.0/24  [10] area: 0.0.0.2
                        directly attached to vlan0
N   192.168.129.0/24  [10] area: 0.0.0.2
                        directly attached to vlan2
N   192.168.130.0/24  [10] area: 0.0.0.2
                        directly attached to vlan3
N   192.168.131.0/24  [10] area: 0.0.0.2
                        directly attached to eth1

===== OSPF router routing table =====
R   172.16.1.1         [10] area: 0.0.0.0, ABR
                        via 172.16.1.9, vlan1
```

Finalment, veure el comportament del BGP i la taula d'encaminament resultant:

```
OpenWrt# show ip bgp
BGP table version is 0, local router ID is 172.16.1.13
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 172.16.1.12/30	0.0.0.0	0		32768	i
*> 192.168.0.0/21	172.16.1.14	0		0 200	i
*> 192.168.64.0/21	172.16.1.9	20		32768	?
*> 192.168.128.0/21	0.0.0.0	0		32768	i

Total number of prefixes 4

Aquí és pot apreciar el diferent comportament de l'atribut Origin i l'impacte sobre el Path. Notar que tot i la ruta a 192.168.64.0/21 és d'una xarxa interna no es coneguda pel BGP amb un origin com a tal. Això és degut a que l'encaminador que l'anuncia no executa BGP i per tant, la ruta s'ha injectat a través del BGP. Com es va comentar en l'apartat corresponent, l'origin és "unknown" i en l'atribut Path apareix com "?". Com a segon exemple, la xarxa

192.168.128.0/21 si es coneguda com a interna, igual que la 192.168.0.0/21, interna de l'AS veí, el 200.

Les rutes queden de la següent forma:

```
OpenWrt# show ip route
Codes: K - kernel route, C - connected, S - static, R - RIP, O - OSPF,
      I - ISIS, B - BGP, > - selected route, * - FIB route

C>* 127.0.0.0/8 is directly connected, lo
O   172.16.1.8/30 [110/10] is directly connected, vlan1, 4d01h02m
C>* 172.16.1.8/30 is directly connected, vlan1
C>* 192.16.1.12/30 is directly connected, vlan4
B>* 192.168.0.0/21 [20/0] via 172.16.1.14, vlan4, 3d23h25m
O>* 192.168.64.0/21 [110/20] via 172.16.1.9, vlan1, 4d01h00m
C>* 192.168.100.0/29 is directly connected, vlan1
O   192.168.128.0/24 [110/10] is directly connected, vlan0, 4d01h02m
C>* 192.168.128.0/24 is directly connected, vlan0
O   192.168.129.0/24 [110/10] is directly connected, vlan2, 4d01h02m
C>* 192.168.129.0/24 is directly connected, vlan2
O   192.168.130.0/24 [110/10] is directly connected, vlan3, 4d01h02m
C>* 192.168.130.0/24 is directly connected, vlan3
O   192.168.131.0/24 [110/10] is directly connected, eth1, 4d01h02m
C>* 192.168.131.0/24 is directly connected, eth1
```

Per acabar, veure com el BGP consumeix més recursos que l'OSPF i per tant, pot ser interessant limitar la seva utilització:

```
Mem: 11272K used, 3104K free, 0K shrd, 212K buff, 3824K cached
Load average: 0.00, 0.00, 0.00 (State: S=sleeping R=running, W=waiting)
```

PID	USER	STATUS	RSS	PPID	%CPU	%MEM	COMMAND
3026	root	R	408	3004	2.8	2.8	top
2763	quagga	S	2620	1	0.0	18.2	bgpd
2750	quagga	S	1384	1	0.0	9.6	ospfd
2738	quagga	S	888	1	0.0	6.1	zebra

A3.2.2. Router 3

En aquest encaminador ja es podrà veure la ruta injectada pel BGP a l'OSPF. Primer es mostra la configuració, on es veu que no s'executa el BGP.

```
Current configuration:
!
hostname OpenWRT@AP3
log stdout
!
debug ospf event
debug ospf packet all
!
interface eth0
!
interface eth1
 ip address 192.168.68.1/24
 ip ospf cost 10
!
interface lo
 ip address 127.0.0.1/8
!
interface vlan0
 ip address 192.168.64.1/24
```

```

!
interface vlan1
 ip address 172.16.1.9/30
 ip ospf cost 10
!
interface vlan2
 ip address 192.168.65.1/24
 ip ospf cost 10
!
interface vlan3
 ip address 192.168.66.1/24
 ip ospf cost 10
!
interface vlan4
 ip address 192.168.67.1/24
 ip ospf cost 10
!
router ospf
 ospf router-id 172.16.1.1
 passive-interface eth1
 network 172.16.1.8/30 area 0.0.0.0
 network 192.168.64.0/24 area 0.0.0.1
 network 192.168.65.0/24 area 0.0.0.1
 network 192.168.66.0/24 area 0.0.0.1
 network 192.168.67.0/24 area 0.0.0.1
 network 192.168.68.0/24 area 0.0.0.1
 area 0.0.0.1 stub no-summary
 area 0.0.0.1 range 192.168.64.0/24 substitute 192.168.64.0/21
 area 0.0.0.1 range 192.168.65.0/24 not-advertise
 area 0.0.0.1 range 192.168.66.0/24 not-advertise
 area 0.0.0.1 range 192.168.67.0/24 not-advertise
 area 0.0.0.1 range 192.168.68.0/24 not-advertise
!
ip forwarding
ipv6 forwarding
!
line vty
!

```

En aquest cas, l'encaminador és ABR, tant de l'area 0.0.0.0 com de la 0.0.0.1:

```

OpenWrt# show ip ospf
 OSPF Routing Process, Router ID: 172.16.1.1
 Supports only single TOS (TOS0) routes
 This implementation conforms to RFC2328
 RFC1583Compatibility flag is disabled
 SPF schedule delay 1 secs, Hold time between two SPF's 1 secs
 Refresh timer 10 secs
 This router is an ABR, ABR type is: Standard (RFC2328)
 Number of external LSA 2. Checksum Sum 0x00018984
 Number of areas attached to this router: 2
 ...

```

Per tant, aquest encaminador sí tindrà constància de la ruta externa que el seu veí ha injectat dins l'OSPF (192.168.0.0/21, sumarytzada):

```

OpenWrt# show ip ospf database

 OSPF Router with ID (172.16.1.1)

 Router Link States (Area 0.0.0.0)

 Link ID          ADV Router      Age Seq#          CkSum Link count
 172.16.1.1       172.16.1.1     1686 0x800000c8 0x086d 1
 172.16.3.1       172.16.3.1     446 0x800000c9 0xf577 1

```

```

Net Link States (Area 0.0.0.0)

Link ID          ADV Router      Age  Seq#           CkSum
172.16.1.10     172.16.3.1     450  0x800000c3    0xf49d

Summary Link States (Area 0.0.0.0)

Link ID          ADV Router      Age  Seq#           CkSum  Route
192.168.64.0    172.16.1.1     657  0x800000c2    0x151b  192.168.64.0/21
192.168.128.0   172.16.3.1     1185 0x800000c2    0x44a9  192.168.128.0/21

Router Link States (Area 0.0.0.1 [Stub])

Link ID          ADV Router      Age  Seq#           CkSum  Link count
172.16.1.1      172.16.1.1     80   0x800000c6    0xcb78  5

Summary Link States (Area 0.0.0.1 [Stub])

Link ID          ADV Router      Age  Seq#           CkSum  Route
0.0.0.0         172.16.1.1     177  0x800000c3    0x6e6e  0.0.0.0/0
172.16.1.8      172.16.1.1     1127 0x800000c1    0xd939  172.16.1.8/30
192.168.128.0   172.16.1.1     36   0x800000c3    0xd214  192.168.128.0/21

AS External Link States

Link ID          ADV Router      Age  Seq#           CkSum  Route
0.0.0.0         172.16.3.1     625  0x800000c2    0xde53  E2 0.0.0.0/0 [0x0]
192.168.0.0     172.16.3.1     1135 0x800000bf    0xab31  E2 192.168.0.0/21 [0x0]

```

Com a conseqüència, aquesta ruta 192.168.0.0 externa ja es coneguda per l'OSPF i per tant encaminable. Qualsevol altre encaminador dins l'àrea, i al tractar-se d'una area totalment stub en termes OSPF, ja no en tindrà constància i directament encaminarà el tràfic cap aquest node, l'ABR:

```

OpenWrt# show ip route ospf
Codes: K - kernel route, C - connected, S - static, R - RIP, O - OSPF,
       I - ISIS, B - BGP, > - selected route, * - FIB route

O>* 0.0.0.0/0 [110/30] via 172.16.1.10, vlan1, 4d01h09m
O 172.16.1.8/30 [110/10] is directly connected, vlan1, 4d01h09m
O>* 192.168.0.0/21 [110/20] via 172.16.1.10, vlan1, 3d23h34m
O 192.168.64.0/24 [110/10] is directly connected, vlan0, 4d01h09m
O 192.168.65.0/24 [110/10] is directly connected, vlan2, 4d01h09m
O 192.168.66.0/24 [110/10] is directly connected, vlan3, 4d01h09m
O 192.168.67.0/24 [110/10] is directly connected, vlan4, 4d01h09m
O 192.168.68.0/24 [110/10] is directly connected, eth1, 4d01h09m
O>* 192.168.128.0/21 [110/20] via 172.16.1.10, vlan1, 4d01h09m

```

La segona ruta externa és la ruta per defecte, que si anunciarà als seus veïns OSPF, que basaran tot l'encaminament en aquesta. Finalment, els recursos consumits són menors, al no executar el BGP:

```

Mem: 11172K used, 3204K free, 0K shrd, 828K buff, 4536K cached
Load average: 0.00, 0.00, 0.00 (State: S=sleeping R=running, W=waiting)

```

```

PID USER      STATUS  RSS  PPID  %CPU  %MEM  COMMAND
17444 root      R       408  17416  2.8   2.8   top
6931 quagga    S      1412    1    0.0   9.8   ospfd
6919 quagga    S       852    1    0.0   5.9   zebra
17415 root      S       636   505   0.0   4.4   dropbear

```

A3.2.3. Router 2

Per acabar amb aquest escenari, es mostra i comenta la configuració del node de l'AS 200. En aquesta ocasió, es pot observar com es controla per mitjà de filtres les rutes que el BGP anuncia o aprèn dels seus veïns. En concret, la configuració aquí detallada permet únicament l'anunci de les xarxes client de cada AS.

```
OpenWrt# show running-config
Building configuration...

Current configuration:
!
hostname OpenWRT@AP2
log stdout
!
interface eth0
!
interface eth1
 ip address 192.168.4.1/24
 ip ospf cost 10
!
interface lo
 ip address 127.0.0.1/8
!
interface vlan0
 ip address 192.168.0.1/24
 ip ospf cost 10
!
interface vlan1
 ip address 192.168.1.1/24
 ip ospf cost 10
!
interface vlan2
 ip address 192.168.2.1/24
 ip ospf cost 10
!
interface vlan3
 ip address 192.168.3.1/24
 ip ospf cost 10
!
interface vlan4
 ip address 172.16.1.14/30
!
router bgp 200
  bgp router-id 172.16.1.14
  network 172.16.1.12/30
  network 192.168.0.0/21
  redistribute ospf
  neighbor 172.16.1.13 remote-as 100
  neighbor 172.16.1.13 prefix-list AS100 in
  neighbor 172.16.1.13 prefix-list AS200 out
!
router ospf
 ospf router-id 172.16.0.0
 redistribute bgp
 network 192.168.0.0/24 area 0.0.0.0
 network 192.168.1.0/24 area 0.0.0.0
 network 192.168.2.0/24 area 0.0.0.0
 network 192.168.3.0/24 area 0.0.0.0
 network 192.168.4.0/24 area 0.0.0.0
 network 192.168.5.0/24 area 0.0.0.0
!
ip prefix-list AS100 seq 5 permit 192.168.128.0/21
ip prefix-list AS100 seq 10 permit 192.168.64.0/21
```

```

ip prefix-list AS100 seq 15 deny any
ip prefix-list AS200 seq 5 permit 192.168.0.0/21
ip prefix-list AS200 seq 10 deny any
!
ip forwarding
ipv6 forwarding
!
line vty
!

```

Com es pot observar, en aquest punt de l'escenari és on és limita al BGP a només intercanviar les xarxes client (192.168.X.X) entre els diferents AS, tal i com s'ha comentat en l'apartat. En la configuració del BGP si es permeten les xarxes troncales, ja que és necessari que el protocol les conegui com a internes.

```

OpenWrt# show ip bgp
BGP table version is 0, local router ID is 172.16.1.14
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 172.16.1.12/30	0.0.0.0	0		32768	i
*> 192.168.0.0/21	0.0.0.0	0		32768	i
*> 192.168.64.0/21	172.16.1.13	20		0 100	?
*> 192.168.128.0/21	172.16.1.13	0		0 100	i

Aquí també es poden observar les diferències entre les dues xarxes accessibles a través del veí AS100. Tanmateix, l'encaminador enrutarà els paquets indistintament i sense diferència aparent. En cas d'existir una altra ruta per a la xarxa 192.168.64.0/21 marcada amb el Path "100 i" s'utilitzaria preferentment. Aquesta situació només es donarà si dins l'AS 100 s'executa el iBGP.

Les diferents taules d'encaminament són les següents:

```

OpenWrt# show ip route ospf
Codes: K - kernel route, C - connected, S - static, R - RIP, O - OSPF,
       I - ISIS, B - BGP, > - selected route, * - FIB route

```

```

O 192.168.0.0/24 [110/10] is directly connected, vlan0, 3d23h43m
O 192.168.1.0/24 [110/10] is directly connected, vlan1, 3d23h43m
O 192.168.2.0/24 [110/10] is directly connected, vlan2, 3d23h43m
O 192.168.3.0/24 [110/10] is directly connected, vlan3, 3d23h43m
O 192.168.4.0/24 [110/10] is directly connected, eth1, 3d23h43m

```

```

OpenWrt# show ip route bgp
Codes: K - kernel route, C - connected, S - static, R - RIP, O - OSPF,
       I - ISIS, B - BGP, > - selected route, * - FIB route

```

```

B>* 192.168.64.0/21 [20/20] via 172.16.1.13, vlan4, 3d23h43m
B>* 192.168.128.0/21 [20/0] via 172.16.1.13, vlan4, 3d23h43m

```

Per acabar, només comentar que la quantitat de recursos utilitzats és semblant a la de l'altre encaminador que executa BGP dins l'escenari, encara que un xic menor donada la baixa carga de l'OSPF.

A3.3. Escenari 2

Durant aquest segon escenari de BGP s'ha volgut estudiar el diferent comportament del BGP com a protocol interior (iBGP) o exterior (eBGP). A més, es interessant començar a veure l'efecte que tenen algun dels atributs del protocol dins d'aquest escenari.

A títol recordatori, l'arquitectura inicial de l'escenari és la següent:

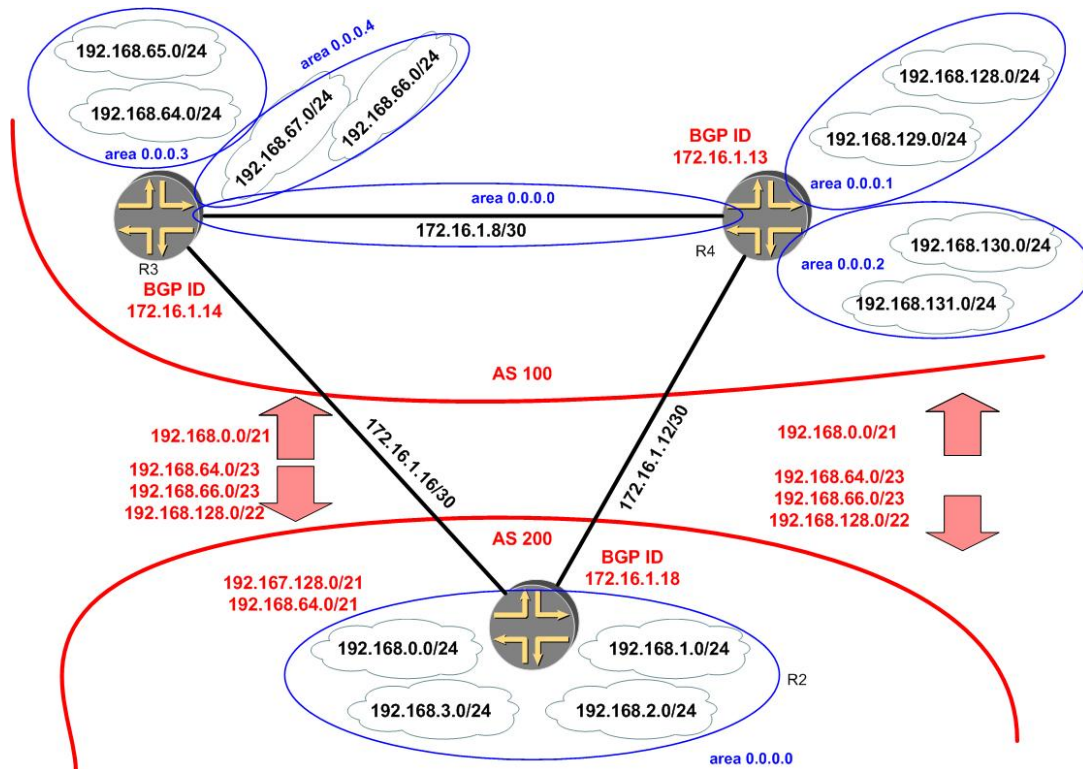


Fig. A3.2 Segon escenari de BGP

A3.3.1. Router 4

La configuració inicial d'aquest equip és la següent. Notar com els veïns BGP s'indiquen a partir d'una IP directament connectada. Aquesta, no té perquè ser estrictament l'adreça de l'enllaç punt a punt entre els veïns, encara que simplifica la gestió.

```
hostname OpenWRT@AP4
log stdout
!
password ap4
enable password ap4
!
interface eth0
!
interface eth1
ip address 192.168.131.1/24
ip ospf cost 10
!
interface lo
```

```

ip address 127.0.0.1/8
!
interface vlan0
ip address 192.168.128.1/24
ip ospf cost 10
!
interface vlan1
ip address 172.16.1.10/30
ip ospf cost 10
!
interface vlan2
ip address 192.168.130.1/24
ip ospf cost 10
!
interface vlan3
ip address 192.168.129.1/24
ip ospf cost 10
!
interface vlan4
ip address 172.16.1.13/30
!
router bgp 100
  bgp router-id 172.16.1.13
  network 192.168.128.0/23
  network 192.168.130.0/23
  aggregate-address 192.168.128.0/22 as-set summary-only
  neighbor 172.16.1.9 remote-as 100
  neighbor 172.16.1.14 remote-as 200
!
router ospf
  ospf router-id 172.16.3.1
  redistribute connected
  passive-interface eth1
  network 172.16.1.8/30 area 0.0.0.0
  network 192.168.128.0/24 area 0.0.0.1
  network 192.168.129.0/24 area 0.0.0.1
  network 192.168.130.0/24 area 0.0.0.2
  network 192.168.131.0/24 area 0.0.0.2
  area 0.0.0.1 range 192.168.128.0/24 substitute 192.168.128.0/23
  area 0.0.0.1 range 192.168.129.0/24 not-advertise
  area 0.0.0.2 range 192.168.130.0/24 substitute 192.168.130.0/23
  area 0.0.0.2 range 192.168.131.0/24 not-advertise
!
ip forwarding
ipv6 forwarding
!
line vty
!

```

Un punt important a l'hora de configurar els veïns BGP és primer especificar els del propi AS i posteriorment els de xarxes remotes. En aquest cas, s'ha configurat un sumari que inclou els dos rangs ja agregats de l'OSPF. S'ha comprovat amb aquesta configuració que les xarxes amb una màscara /23 no són anunciades.

Un cop la xarxa esta configurada el node estableix una sessió BGP amb els dos veïns possibles, un d'intern i l'altre extern:

```

BGP neighbor is 172.16.1.9, remote AS 100, local AS 100, internal link
  BGP version 4, remote router ID 172.16.1.17
  BGP state = Established, up for 01:13:03
  Last read 00:00:02, hold time is 180, keepalive interval is 60 seconds
  Neighbor capabilities:
    Route refresh: advertised and received(old & new)
    Address family IPv4 Unicast: advertised and received

```

```

Received 77 messages, 0 notifications, 0 in queue
Sent 78 messages, 0 notifications, 0 in queue
Route refresh request: received 0, sent 0
Minimum time between advertisement runs is 5 seconds

```

```

For address family: IPv4 Unicast
Community attribute sent to this neighbor(both)
4 accepted prefixes

```

```

Connections established 1; dropped 0
Last reset never
Local host: 172.16.1.10, Local port: 3483
Foreign host: 172.16.1.9, Foreign port: 179
Nextthop: 172.16.1.10
Nextthop global: ::
Nextthop local: ::
BGP connection: non shared network
Read thread: on Write thread: off

```

```

BGP neighbor is 172.16.1.14, remote AS 200, local AS 100, external link
BGP version 4, remote router ID 172.16.1.18
BGP state = Established, up for 00:20:42
Last read 00:00:42, hold time is 180, keepalive interval is 60 seconds
Neighbor capabilities:
Route refresh: advertised and received(old & new)
Address family IPv4 Unicast: advertised and received
Received 23 messages, 0 notifications, 0 in queue
Sent 25 messages, 0 notifications, 0 in queue
Route refresh request: received 0, sent 0
Minimum time between advertisement runs is 30 seconds

```

```

For address family: IPv4 Unicast
Community attribute sent to this neighbor(both)
2 accepted prefixes

```

```

Connections established 1; dropped 0
Last reset never
Local host: 172.16.1.13, Local port: 3484
Foreign host: 172.16.1.14, Foreign port: 179
Nextthop: 172.16.1.13
Nextthop global: ::
Nextthop local: ::
BGP connection: non shared network
Read thread: on Write thread: off

```

Amb això, el protocol BGP obté la base de dades de camins. Notar que a diferència de l'escenari anterior no apareixen xarxes amb l'atribut "origin" amb valor "incomplete". En aquest cas, totes les xarxes són externes o internes.

```

BGP table version is 0, local router ID is 172.16.1.13
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

```

Network	Next Hop	Metric	LocPrf	Weight	Path
* i192.168.0.0/23	172.16.1.18	0	100	0	200 i
*>	172.16.1.14	0		0	200 i
* i192.168.2.0/23	172.16.1.18	0	100	0	200 i
*>	172.16.1.14	0		0	200 i
*>i192.168.64.0/23	172.16.1.9	0	100	0	i
*>i192.168.66.0/23	172.16.1.9	0	100	0	i
*> 192.168.128.0/22	0.0.0.0			32768	i
s> 192.168.128.0/23	0.0.0.0	0		32768	i
s> 192.168.130.0/23	0.0.0.0	0		32768	i

Com que el BGP només instal·la una única ruta per a cada xarxa cal veure la taula d'encaminament:

```
OpenWrt# show ip route
Codes: K - kernel route, C - connected, S - static, R - RIP, O - OSPF,
       I - ISIS, B - BGP, > - selected route, * - FIB route

C>* 127.0.0.0/8 is directly connected, lo
O   172.16.1.8/30 [110/10] is directly connected, vlan1, 01:37:20
C>* 172.16.1.8/30 is directly connected, vlan1
C>* 172.16.1.12/30 is directly connected, vlan4
B>* 192.168.0.0/23 [20/0] via 172.16.1.14, vlan4, 00:20:37
B>* 192.168.2.0/23 [20/0] via 172.16.1.14, vlan4, 00:20:37
B   192.168.64.0/23 [200/0] via 172.16.1.9, vlan1, 01:12:58
O>* 192.168.64.0/23 [110/20] via 172.16.1.9, vlan1, 01:13:06
B   192.168.66.0/23 [200/0] via 172.16.1.9, vlan1, 01:12:58
O>* 192.168.66.0/23 [110/20] via 172.16.1.9, vlan1, 01:13:06
O   192.168.128.0/24 [110/10] is directly connected, vlan0, 01:37:26
C>* 192.168.128.0/24 is directly connected, vlan0
O   192.168.129.0/24 [110/10] is directly connected, vlan3, 01:37:26
C>* 192.168.129.0/24 is directly connected, vlan3
O   192.168.130.0/24 [110/10] is directly connected, vlan2, 01:37:26
C>* 192.168.130.0/24 is directly connected, vlan2
O   192.168.131.0/24 [110/10] is directly connected, eth1, 01:37:26
C>* 192.168.131.0/24 is directly connected, eth1
```

Un cop vista, és necessari fer dos comentaris al respecte. Primer de tot notar com hi ha una ruta BGP (iBGP) que no s'instal·la dins la taula d'encaminament tot i ser el millor camí segons BGP. Aquest efecte és donat a que també s'està executant l'OSPF com a protocol interior (IGP) i les rutes d'aquest són preferides respecte les de BGP. Notar com realment si existeix una ruta implícita cap a 192.168.66.0/23, però que aquesta és de l'OSPF. Finalment, i com a segona observació notar com el BGP si instal·la correctament les rutes per anar fins a l'AS veí.

A3.3.2. Router 3

A continuació es realitza un anàlisi idèntic que per al router 4. La configuració és la següent:

```
hostname OpenWRT@AP3
log stdout
!
debug ospf event
debug ospf packet all
!
interface eth0
!
interface eth1
 ip address 192.168.67.1/24
 ip ospf cost 10
!
interface lo
 ip address 127.0.0.1/8
!
interface vlan0
 ip address 192.168.64.1/24
 ip ospf cost 10
!
interface vlan1
```

```

ip address 172.16.1.9/30
ip ospf cost 10
!
interface vlan2
ip address 192.168.65.1/24
ip ospf cost 10
!
interface vlan3
ip address 172.16.1.17/30
!
interface vlan4
ip address 192.168.66.1/24
ip ospf cost 10
!
router bgp 100
  bgp router-id 172.16.1.17
  network 192.168.64.0/23
  network 192.168.66.0/23
  neighbor 172.16.1.10 remote-as 100
  neighbor 172.16.1.18 remote-as 200
!
router ospf
  ospf router-id 172.16.1.1
  passive-interface eth1
  network 172.16.1.8/30 area 0.0.0.0
  network 192.168.64.0/24 area 0.0.0.3
  network 192.168.65.0/24 area 0.0.0.3
  network 192.168.66.0/24 area 0.0.0.4
  network 192.168.67.0/24 area 0.0.0.4
  area 0.0.0.3 stub no-summary
  area 0.0.0.3 range 192.168.64.0/24 substitute 192.168.64.0/23
  area 0.0.0.3 range 192.168.65.0/24 not-advertise
  area 0.0.0.4 stub
  area 0.0.0.4 range 192.168.66.0/24 substitute 192.168.66.0/23
  area 0.0.0.4 range 192.168.67.0/24 not-advertise
!
ip forwarding
ipv6 forwarding
!
line vty
!

```

Un cop la sessió BGP s'ha establert correctament es pot veure com la base de dades BGP és tant extensa com en el seu veí intern i com les rutes instal·lades són vàlides.

```

penWrt# show ip bgp
BGP table version is 0, local router ID is 172.16.1.17
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 192.168.0.0/23	172.16.1.18	0		0	200 i
* i	172.16.1.14	0	100	0	200 i
*> 192.168.2.0/23	172.16.1.18	0		0	200 i
* i	172.16.1.14	0	100	0	200 i
*> 192.168.64.0/23	0.0.0.0	0		32768	i
*> 192.168.66.0/23	0.0.0.0	0		32768	i
*>i192.168.128.0/22	172.16.1.10		100	0	i

Finalment, la taula d'encaminament segueix el camí òptim per a cada xarxa anunciada:

```

OpenWrt# show ip route
Codes: K - kernel route, C - connected, S - static, R - RIP, O - OSPF,
       I - ISIS, B - BGP, > - selected route, * - FIB route

C>* 127.0.0.0/8 is directly connected, lo
O   172.16.1.8/30 [110/10] is directly connected, vlan1, 01:26:13
C>* 172.16.1.8/30 is directly connected, vlan1
O>* 172.16.1.12/30 [110/20] via 172.16.1.10, vlan1, 01:11:50
C>* 172.16.1.16/30 is directly connected, vlan3
B>* 192.168.0.0/23 [20/0] via 172.16.1.18, vlan3, 00:05:34
B>* 192.168.2.0/23 [20/0] via 172.16.1.18, vlan3, 00:05:34
O   192.168.64.0/24 [110/10] is directly connected, vlan0, 01:26:13
C>* 192.168.64.0/24 is directly connected, vlan0
O   192.168.65.0/24 [110/10] is directly connected, vlan2, 01:26:13
C>* 192.168.65.0/24 is directly connected, vlan2
O   192.168.66.0/24 [110/10] is directly connected, vlan4, 01:26:13
C>* 192.168.66.0/24 is directly connected, vlan4
O   192.168.67.0/24 [110/10] is directly connected, eth1, 01:26:13
C>* 192.168.67.0/24 is directly connected, eth1
B>* 192.168.128.0/22 [200/0] via 172.16.1.10, vlan1, 01:11:48
O>* 192.168.128.0/23 [110/20] via 172.16.1.10, vlan1, 01:11:51
O>* 192.168.130.0/23 [110/20] via 172.16.1.10, vlan1, 01:11:51

```

A3.3.3. Router 2

Per acabar amb l'escenari segon es mostra únicament la configuració i la taula d'encaminament generada, proves del correcte funcionament de l'escenari. La configuració és la detallada a continuació.

```

hostname OpenWRT@AP2
log stdout
!
interface eth0
!
interface eth1
 ip address 192.168.3.1/24
 ip ospf cost 10
!
interface lo
 ip address 127.0.0.1/8
!
interface vlan0
 ip address 192.168.1.1/24
 ip ospf cost 10
!
interface vlan1
 ip address 192.168.0.1/24
 ip ospf cost 10
!
interface vlan2
 ip address 192.168.2.1/24
 ip ospf cost 10
!
interface vlan3
 ip address 172.16.1.18/30
!
interface vlan4
 ip address 172.16.1.14/30
!
router bgp 200
 bgp router-id 172.16.1.18
 bgp default local-preference 50
 network 192.168.0.0/23
 network 192.168.2.0/23

```

```

neighbor 172.16.1.13 remote-as 100
neighbor 172.16.1.17 remote-as 100
!
router ospf
  ospf router-id 192.168.1.1
  redistribute bgp
  network 0.0.0.0/0 area 0.0.0.0
!
ip forwarding
ipv6 forwarding

```

I les rutes resultants son:

```

Total number of prefixes 5
OpenWrt# show ip route
Codes: K - kernel route, C - connected, S - static, R - RIP, O - OSPF,
       I - ISIS, B - BGP, > - selected route, * - FIB route

C>* 127.0.0.0/8 is directly connected, lo
O   172.16.1.12/30 [110/10] is directly connected, vlan4, 00:24:01
C>* 172.16.1.12/30 is directly connected, vlan4
O   172.16.1.16/30 [110/10] is directly connected, vlan3, 00:24:01
C>* 172.16.1.16/30 is directly connected, vlan3
O   192.168.0.0/24 [110/10] is directly connected, vlan1, 00:24:01
C>* 192.168.0.0/24 is directly connected, vlan1
O   192.168.1.0/24 [110/10] is directly connected, vlan0, 00:24:01
C>* 192.168.1.0/24 is directly connected, vlan0
O   192.168.2.0/24 [110/10] is directly connected, vlan2, 00:24:01
C>* 192.168.2.0/24 is directly connected, vlan2
O   192.168.3.0/24 [110/10] is directly connected, eth1, 00:24:01
C>* 192.168.3.0/24 is directly connected, eth1
B>* 192.168.64.0/23 [20/0] via 172.16.1.13, vlan4, 00:21:35
B>* 192.168.66.0/23 [20/0] via 172.16.1.13, vlan4, 00:21:35
B>* 192.168.128.0/22 [20/0] via 172.16.1.13, vlan4, 00:21:35

```

Notar com si existeix un únic punt d'entrada de l'AS 200 cap al 100. Aquest és el comportament habitual del BGP. En el sentit contrari, tal i com s'ha mostrat en router 4 s'ha alterat aquest comportament per a que existeixin dos punts d'entrada de l'AS 100 cap al 200. Aquestes rutes externes s'inclouen dins la base de dades l'OSPF:

```

OpenWrt# show ip ospf database

      OSPF Router with ID (192.168.1.1)

          Router Link States (Area 0.0.0.0)

Link ID        ADV Router    Age Seq#           CkSum  Link count
192.168.1.1    192.168.1.1   1407 0x80000003 0x1103 7

          AS External Link States

Link ID        ADV Router    Age Seq#           CkSum  Route
192.168.64.0   192.168.1.1   171 0x80000002 0x07d7 E2 192.168.64.0/23 [0x0]
192.168.66.0   192.168.1.1   1303 0x80000001 0xf2ea E2 192.168.66.0/23 [0x0]
192.168.128.0  192.168.1.1   1303 0x80000001 0x3c65 E2 192.168.128.0/22 [0x0]

```

Finalment, com a mostra que el BGP consumeix més recursos que l'OSPF comentar que els requeriments de memòria per aquest escenari son del 14%, front el 5% de OSPF. Per tant, cal acotar quins son els nodes que executen el BGP i dotar-los d'una configuració OSPF mínima.

A3.4. Escenari 3

Conegudes les capacitats del protocol BGP com a encaminador interior i exterior així com també alguns dels atributs de mètrica, aquest tercer escenari vol estudiar l'ús de l'atribut de comunitat entre d'altres. La funció principal és actuar de filtre sobre els prefixos que s'anuncien o es reben aplicant-los diverses polítiques posteriors.

L'arquitectura utilitzada és a següent:

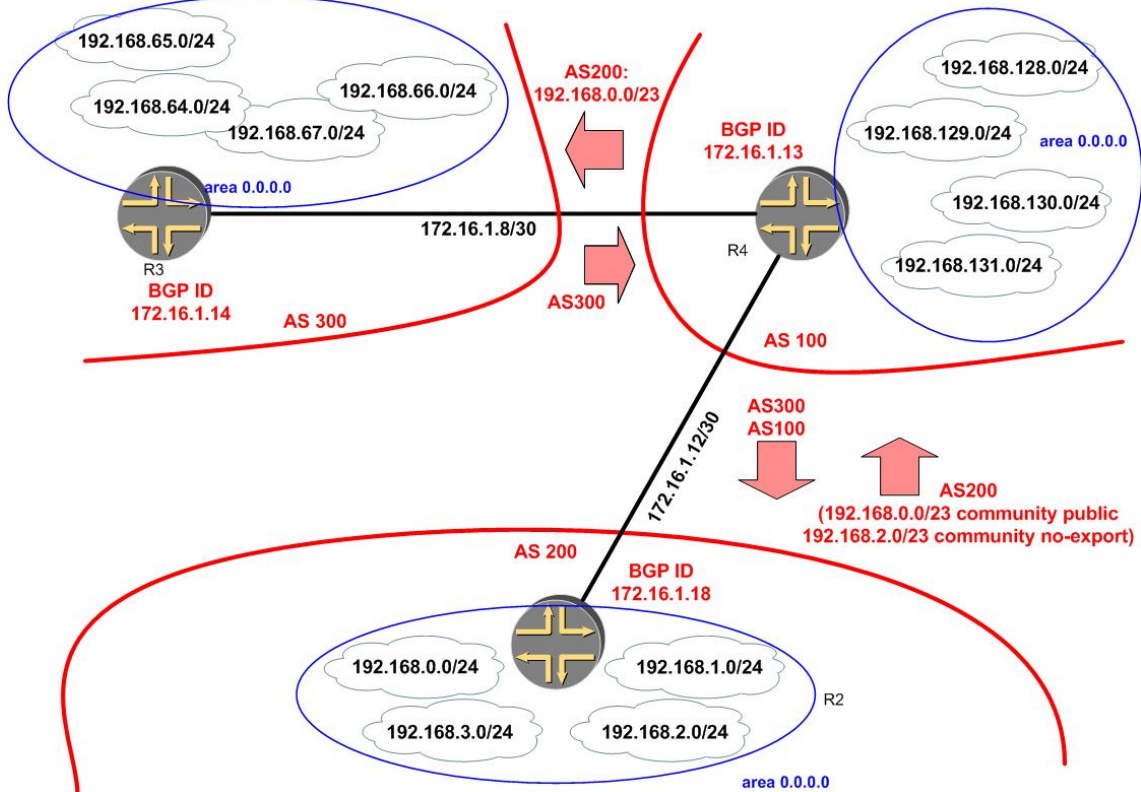


Fig. A3.3 Una de les maquetes del tercer escenari

Segons es veu en la figura, l'AS200 no sumaritza completament totes les seves adreces per diferenciar la visibilitat amb l'AS300. Per a fer-ho, l'AS200 envia les dues xarxes amb atributs ben diferents. La primera, 192.168.0.0/23, està marcada pública (o internet) i per tant, serà advertida a tots els seus veïns. La segona xarxa anunciada cap a l'AS100, la 192.168.2.0/23, està marcada com a no-export. Com a conseqüència, aquesta xarxa no serà advertida cap a l'AS300. Per acabar de comentar l'escenari, l'AS100 anunciarà les seves xarxes a tots els veïns i l'AS300 també.

En aquesta ocasió, la configuració comentada que es detalla correspon a la prova explicada donades les poques diferències amb la configuració inicial de l'escenari.

A3.4.1. Router 4 (AS100)

La configuració del router i els diferents protocols és la següent:

```

interface eth0
!
interface eth1
 ip address 192.168.131.1/24
 ip ospf cost 10
!
interface lo
 ip address 127.0.0.1/8
!
interface vlan0
 ip address 192.168.128.1/24
 ip ospf cost 10
!
interface vlan1
 ip address 172.16.1.10/30
 ip ospf cost 10
!
interface vlan2
 ip address 192.168.130.1/24
 ip ospf cost 10
!
interface vlan3
 ip address 192.168.129.1/24
 ip ospf cost 10
!
interface vlan4
 ip address 172.16.1.13/30
!
router bgp 100
  bgp router-id 172.16.1.13
  network 192.168.128.0/23
  network 192.168.130.0/23
  aggregate-address 192.168.128.0/21 as-set summary-only
  neighbor 172.16.1.9 remote-as 300
  neighbor 172.16.1.14 remote-as 200
!
router ospf
  ospf router-id 172.16.3.1
  redistribute bgp
  passive-interface eth1
  area 0.0.0.0 range 0.0.0.0/0
!
ip route 192.168.128.0/21 Null0
!
ip forwarding
ipv6 forwarding

```

Notar la ruta estàtica cap a 192.168.128.0/21 via Null0. Aquesta ruta indica al router que descarti els paquets que no tinguin una entrada vàlida en la seva taula d'encaminament. Donat que com a node BGP anuncia tot el rang és una bona pràctica que també sigui el normalment anomenat blackhole o forat negre.

La llista de xarxes presents en la base de dades és completa, ja que els dos prefixos anunciats per l'AS200 hi són presents. Notar la presència de les dues xarxes suprimides donada l'agregació de sumaris:

```

OpenWrt# show ip bgp
BGP table version is 0, local router ID is 172.16.1.13
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal

```

Origin codes: i - IGP, e - EGP, ? - incomplete

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 192.168.0.0/23	172.16.1.14	0		0	200 i
*> 192.168.2.0/23	172.16.1.14	0		0	200 i
*> 192.168.64.0/21	172.16.1.9	0		0	300 i
*> 192.168.128.0/21	0.0.0.0			32768	i
s> 192.168.128.0/23	0.0.0.0	0		32768	i
s> 192.168.130.0/23	0.0.0.0	0		32768	i

Per tant, la taula d'encaminament és la següent:

```
OpenWrt# show ip route
Codes: K - kernel route, C - connected, S - static, R - RIP, O - OSPF,
       I - ISIS, B - BGP, > - selected route, * - FIB route

C>* 127.0.0.0/8 is directly connected, lo
C>* 172.16.1.8/30 is directly connected, vlan1
C>* 172.16.1.12/30 is directly connected, vlan4
B>* 192.168.0.0/23 [20/0] via 172.16.1.14, vlan4, 00:11:25
B>* 192.168.2.0/23 [20/0] via 172.16.1.14, vlan4, 00:11:25
B>* 192.168.64.0/21 [20/0] via 172.16.1.9, vlan1, 00:11:26
S>* 192.168.128.0/21 [1/0] is directly connected, Null0, bh
C>* 192.168.128.0/24 is directly connected, vlan0
C>* 192.168.129.0/24 is directly connected, vlan3
C>* 192.168.130.0/24 is directly connected, vlan2
C>* 192.168.131.0/24 is directly connected, eth1
```

Un cop més, veure quines seran les rutes que el BGP introduirà dins de l'IGP (OSPF en aquest cas):

AS External Link States

Link ID	ADV Router	Age	Seq#	CkSum	Route
192.168.0.0	172.16.3.1	707	0x80000001	0x474e E2	192.168.0.0/23 [0x0]
192.168.2.0	172.16.3.1	707	0x80000001	0x3162 E2	192.168.2.0/23 [0x0]
192.168.64.0	172.16.3.1	708	0x80000001	0x66f4 E2	192.168.64.0/21 [0x0]

Com era d'esperar, totes les xarxes del eBGP s'han afegit dins la base de dades. Així doncs, aquest node l'anunciarà cap a dins de l'OSPF i enviat per tota la xarxa interna amb el correcte punt de sortida cap a la xarxa. Finalment, apreciar que el consum de recursos és elevat.

Mem: 12156K used, 2220K free, 0K shrd, 128K buff, 2732K cached
 Load average: 0.00, 0.00, 0.00 (State: S=sleeping R=running, W=waiting)

PID	USER	STATUS	RSS	PPID	%CPU	%MEM	COMMAND
7049	root	R	408	6823	11.5	2.8	top
6777	root	S	2976	6512	0.0	17.7	vttysh
7006	quagga	S	2380	1	0.0	16.5	bgpd
6993	quagga	S	1156	1	0.0	8.0	ospfd
6981	quagga	S	852	1	0.0	5.9	zebra

A3.4.2. Router 3 (AS300)

La configuració de l'AS300, que anuncia totes les seves xarxes públiques, és:

```
debug ospf event
debug ospf packet all
!
interface eth0
!
interface eth1
 ip address 192.168.67.1/24
 ip ospf cost 10
!
interface lo
 ip address 127.0.0.1/8
!
interface vlan0
 ip address 192.168.64.1/24
 ip ospf cost 10
!
interface vlan1
 ip address 172.16.1.9/30
 ip ospf cost 10
!
interface vlan2
 ip address 192.168.65.1/24
 ip ospf cost 10
!
interface vlan3
 ip address 172.16.1.17/30
 shutdown
!
interface vlan4
 ip address 192.168.66.1/24
 ip ospf cost 10
!
router bgp 300
 bgp router-id 172.16.1.17
 network 192.168.64.0/21
 neighbor 172.16.1.10 remote-as 100
 neighbor 172.16.1.10 soft-reconfiguration inbound
 neighbor 172.16.1.10 route-map AS300 out
!
router ospf
 ospf router-id 172.16.1.1
 redistribute bgp
 passive-interface eth1
!
ip route 192.168.64.0/21 Null0
!
ip prefix-list NET seq 5 permit 192.168.64.0/21
ip prefix-list NET seq 10 deny any
!
route-map AS300 permit 5
 match ip address prefix-list NET
 set community internet
!

ip forwarding
ipv6 forwarding
!
line vty
!
```


El resultat esperat és la visibilitat de totes les xarxes. És a dir, que l'AS100 accepta en transit de l'AS100 cap al 200.

```
OpenWrt# show ip bgp
BGP table version is 0, local router ID is 172.16.1.17
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete
```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 192.168.0.0/23	172.16.1.10			0	100 200 i
*> 192.168.64.0/21	0.0.0.0	0		32768	i
*> 192.168.128.0/21	172.16.1.10			0	100 i

En consum de recursos es manté elevat, però dins de les cotes ja observades per al BGP.

```
Mem: 12056K used, 2320K free, 0K shrd, 124K buff, 2456K cached
Load average: 0.00, 0.00, 0.00 (State: S=sleeping R=running, W=waiting)
  PID USER      STATUS  RSS  PPID  %CPU  %MEM  COMMAND
 10847 root        R       408 10642 11.4   2.8  top
 10534 root        S      2716  9738  0.0  18.8  vtysh
   872 quagga     S      2244    1  0.0  15.6  bgpd
 10755 quagga     S      1092    1  0.0   7.5  ospfd
 10743 quagga     S       812    1  0.0   5.6  zebra
```

A3.4.3. Router 2 (AS200)

La configuració de l'AS200, que anuncia totes les seves xarxes publiques:

```
log stdout
!
interface eth0
!
interface eth1
 ip address 192.168.3.1/24
 ip ospf cost 10
!
interface lo
 ip address 127.0.0.1/8
!
interface vlan0
 ip address 192.168.1.1/24
 ip ospf cost 10
!
interface vlan1
 ip address 192.168.0.1/24
 ip ospf cost 10
!
interface vlan2
 ip address 192.168.2.1/24
 ip ospf cost 10
!
interface vlan3
 ip address 172.16.1.18/30
!
interface vlan4
 ip address 172.16.1.14/30
!
router bgp 200
 bgp router-id 172.16.1.18
 bgp default local-preference 50
 network 192.168.0.0/23
```

```

network 192.168.2.0/23
neighbor 172.16.1.13 remote-as 100
neighbor 172.16.1.13 route-map AS200 out
!
router ospf
ospf router-id 192.168.1.1
redistribute bgp
network 0.0.0.0/0 area 0.0.0.0
!
ip route 192.168.0.0/21 Null0
!
ip prefix-list NET1 seq 5 permit 192.168.0.0/23
ip prefix-list NET1 seq 10 deny any
ip prefix-list NET2 seq 5 permit 192.168.2.0/23
ip prefix-list NET2 seq 10 deny any
!
route-map AS200 permit 5
match ip address prefix-list NET1
set community internet
!
route-map AS200 permit 20
match ip address prefix-list NET2
set community no-export
!
ip forwarding
ipv6 forwarding
!
line vty
!

```

La base de dades de BGP resultant és senzilla i mostra les rutes de l'AS100 i 300, tal i com era d'esperar:

```

OpenWrt# show ip bgp
BGP table version is 0, local router ID is 172.16.1.18
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal
Origin codes: i - IGP, e - EGP, ? - incomplete

```

Network	Next Hop	Metric	LocPrf	Weight	Path
*> 192.168.0.0/23	0.0.0.0	0		32768	i
*> 192.168.2.0/23	0.0.0.0	0		32768	i
*> 192.168.64.0/21	172.16.1.13			0 100 300	i
*> 192.168.128.0/21	172.16.1.13			0 100	i

La taula d'encaminament reflecteix la base dades i per tant, totes les xarxes de l'escenari s'encaminen.

```

OpenWrt# show ip route
Codes: K - kernel route, C - connected, S - static, R - RIP, O - OSPF,
I - ISIS, B - BGP, > - selected route, * - FIB route

```

```

C>* 127.0.0.0/8 is directly connected, lo
O 172.16.1.12/30 [110/10] is directly connected, vlan4, 22:44:31
C>* 172.16.1.12/30 is directly connected, vlan4
O 172.16.1.16/30 [110/10] is directly connected, vlan3, 22:44:31
C>* 172.16.1.16/30 is directly connected, vlan3
S>* 192.168.0.0/21 [1/0] is directly connected, Null0, bh
O 192.168.0.0/24 [110/10] is directly connected, vlan1, 22:44:31
C>* 192.168.0.0/24 is directly connected, vlan1
O 192.168.1.0/24 [110/10] is directly connected, vlan0, 22:44:31
C>* 192.168.1.0/24 is directly connected, vlan0
O 192.168.2.0/24 [110/10] is directly connected, vlan2, 22:44:31
C>* 192.168.2.0/24 is directly connected, vlan2
O 192.168.3.0/24 [110/10] is directly connected, eth1, 22:44:31
C>* 192.168.3.0/24 is directly connected, eth1
B>* 192.168.64.0/21 [20/0] via 172.16.1.13, vlan4, 00:09:58
B>* 192.168.128.0/21 [20/0] via 172.16.1.13, vlan4, 00:09:58

```

Com en el cas anterior, es mostra el consum de recursos, que és manté elevat.

```
Mem: 11024K used, 3352K free, 0K shrd, 300K buff, 3480K cached  
Load average: 0.00, 0.00, 0.00 (State: S=sleeping R=running, W=waiting)
```

PID	USER	STATUS	RSS	PPID	%CPU	%MEM	COMMAND
1842	root	R	408	1776	2.9	2.8	top
1775	root	S	624	437	0.9	4.3	dropbear
1810	quagga	S	2660	1	0.0	18.5	bgpd
1242	quagga	S	1240	1	0.0	8.6	ospfd
1239	quagga	S	848	1	0.0	5.8	zebra

ANNEX 4: PLANIFICACIÓ I PRESSUPOST DEL PROJECTE

A4.1. Planificació

El desenvolupament d'aquest estudi està clarament diferenciat en tres etapes. Durant la primera etapa es realitza una presa de contacte amb Guifi.net i l'entorn de treball a utilitzar en quan a hardware es refereix i distribucions Linux a implantar-hi. Durant aquest primer període s'han de redactar una llista de propostes aplicables a Guifi.net i que es consensuaran per veure el seu interès i viabilitat.

La segona etapa, que s'inicia després de veure quins són els objectius clars del projecte, correspon a una presa de contacte amb les tecnologies concretes que s'utilitzaran. Aquestes tasques s'han de realitzar amb tot l'equipament preparat i per tant ja es necessari disposar dels WRT54GL operatius i amb tot el programari disponible.

Finalment, la tercera etapa correspon a la realització de la o les propostes plantejades i posterior extracció de conclusions i quelcom d'utilitat per a Guifi.net.

Per a la redacció del document no es destina un gaire espai temporal dedicat, sinó que la redacció es durà a terme a mesura que la realització del projecte ho permeti. També es considera la realització de reunions setmanals amb el tutor per tal d'analitzar la realització del projecte, definició de tests i objectius, etc

El diagrama de la figura A4.1 mostra una visió més gràfica de la planificació inicial del projecte.

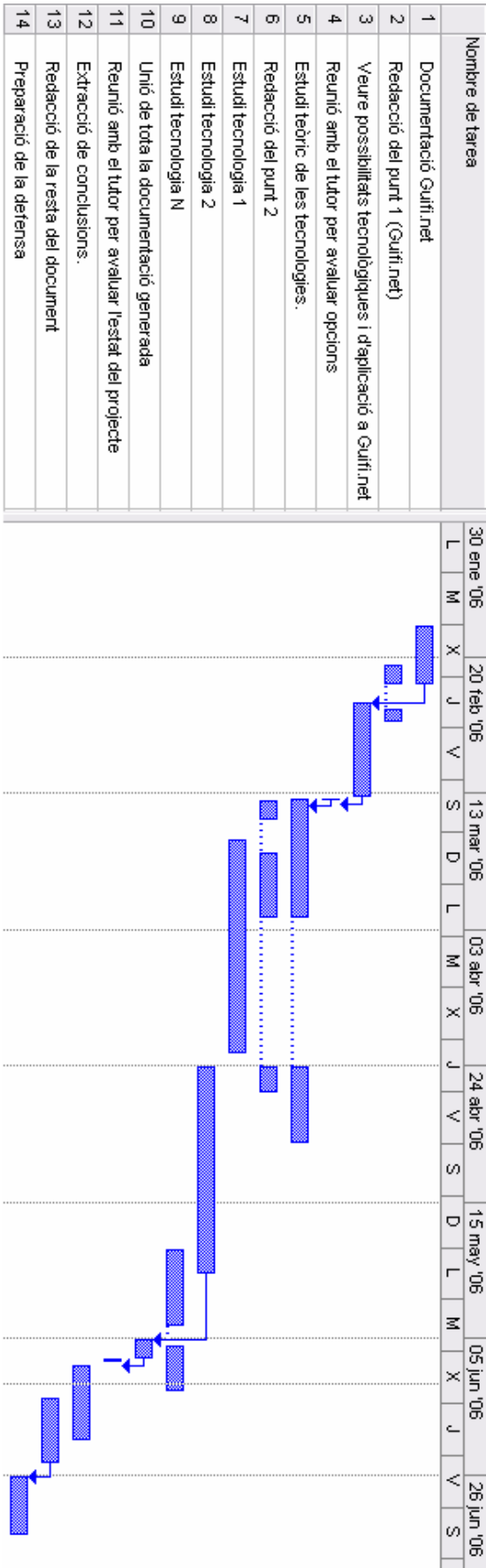


Fig. A4.1 Diagrama de Gannt del projecte

A4.2. Avaluació i pressupost

L'avaluació del desenvolupament d'aquest estudi és positiva. S'ha seguit la planificació segons la definició inicial de la figura A4.1 si s'elimina la tasca addicional (Estudi de la tecnologia N) que es va planificar, i que s'ha canviat per l'estudi d'aplicació a Guifi.net, de durada més breu. Segons la definició inicial, el projecte es va planificar fins el 25 de juny com a data màxima per a la redacció del document. El compliment de la planificació i les modificacions en la normativa vigent han permès acabar aquest estudi deu dies abans de l'estimat.

Per a la realització del treball s'ha considerat una setmana laboral de sis dies, amb una jornada de quatre hores diàries. Per tant, des del 15 de Febrer fins al 15 de Juny es consideren un total de 412 hores.

Sobre el total de 412 hores treballades, es pot obtenir la distribució percentil de cada una de les tasques del projecte. La següent figura mostra com s'ha repartit el recurs temps entre les tasques de documentació, reunions, realització dels escenaris i redacció del document.

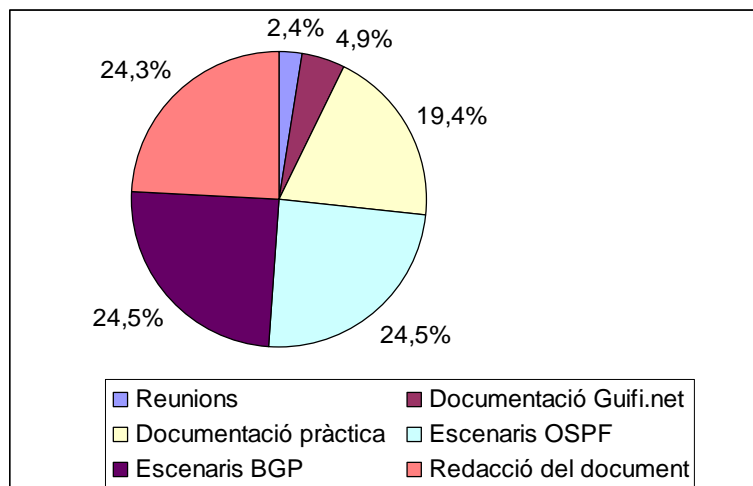


Fig. A4.2 Classificació de les hores invertides en la realització del projecte

A continuació la taula A4.1 mostra un breu pressupost econòmic del projecte.

Material	Unitats	Preu / u.	Preu
Linksys WRT54GL	3	65 €	195 €
Cable UTP 1m.	4	3 €	12 €
Mà d'obra			
Enginyer senior	412	60 €	24.720 €
TOTAL			24.927 €

Taula. A4.1 Pressupost econòmic del projecte

ANNEX 5: GNU Free Documentation License

GNU Free Documentation License
Version 1.2, November 2002

Copyright (C) 2000,2001,2002 Free Software Foundation, Inc.
51 Franklin St, Fifth Floor, Boston, MA 02110-1301 USA
Everyone is permitted to copy and distribute verbatim copies
of this license document, but changing it is not allowed.

0. PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondly, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical

connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other

implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

- A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous version(which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.
- B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.
- C. State on the Title page the name of the publisher of the Modified Version, as the publisher.
- D. Preserve all the copyright notices of the Document.
- E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.
- F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.
- G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.
- H. Include an unaltered copy of this License.
- I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.
- J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.
- K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.
- L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.
- M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.
- N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.
- O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all

of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties--for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements".

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a

copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See <http://www.gnu.org/copyleft/>.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation.