PRACTICAL SECURITY APPROACHES AGAINST BORDER GATEWAY PROTOCOL (BGP) SESSION HIJACKING ATTACKS AND MISCONFIGURATION ERRORS

BY

STEPHEN BRAKO OTI
(B.Sc. INFORMATION TECHNOLOGY)

A THESIS SUBMITTED TO THE DEPARTMENT OF COMPUTER SCIENCE

KWAME NKRUMAH UNIVERSITY OF SCIENCE AND TECHNOLOGY

IN PARTIAL FULFILLMENT OF THE REQUIREMENT FOR THE DEGREE OF MPHIL INFORMATION TECHNOLOGY

©JULY 2012

DECLARATION

This thesis is a presentation of my perspectives on the subject matter and contains to the best of my knowledge; no material published by another for an award. Wherever the contributions of others are involved, every effort is made to indicate this clearly, with due reference to the literature, and acknowledgement of collaborative research and discussions. The work was done under the guidance of Dr. J. B. Hayfron-Acquah and Mr. Stephen Agyapong of the Computer Science Department, KNUST.

STEPHEN BRAKO OTI ……………………… 30TH JULY 2012
(PG5094110) (SIGNATURE) (DATE)

CERTIFIED BY:

J.B. HAYFRON-ACQUAH ……………………… 2/1/13
(SUPERVISOR'S NAME) (SIGNATURE) (DATE)

CERTIFIED BY:

J.B. HAYFRON-ACQUAH ……………………… 2/1/13
(HEAD OF DEPARTMENT) (SIGNATURE) (DATE)

DEDICATION

This work is dedicated to the memory of my late father, KINGSLEY OTI.

To my godfather HON. ISHMAEL TETTEH ARYEETEY without whom I would not have come this far.

ACKNOWLEDGEMENT

I wish to acknowledge my supervisors, Dr. J. B. Hayfron-Acquah and Mr. Stephen Agyapong for their invaluable suggestions, guidance and encouragement through the entire research period. Day and night I bothered these devoted intellectuals yet they always had time for me; and for that I am grateful.

I also wish to acknowledge the invaluable contributions of Gervin Appiah; my work colleague at BBH (Broadband Home Ltd); for his exceptional technical insight and support throughout the entire period. Millicent Otchere kept the research on track as a project manager; always making sure I kept to the schedule and delivered the chapters on time.

Lastly I wish to acknowledge all my friends and siblings who in diverse ways have supported and contributed to the successful completion of this thesis; my heartfelt gratitude goes out to you all.

ABSTRACT

The border gateway protocol (BGP) is the default inter domain routing protocol used on the internet for exchanging information between autonomous systems. Available literature suggests that BGP is vulnerable to session hijacking attacks and misconfiguration errors. There are a number of proposals aimed at improving BGP security which have not been fully implemented. This work examines a number of approaches proposed for the security of BGP through a comparative study and identifies the reasons why these proposals have not been implemented on a commercial scale. This work also analyses the architecture of internet routing and the design of BGP while focusing on the problem of BGP misconfiguration and session hijacking attacks. Secondary data from the Route-Views servers are used as a baseline for the quantitative study of BGP misconfiguration. Using Graphical Network Simulator 3 (GNS-3), a session hijack is demonstrated and a solution which involves the implementation of route filtering, policy-maps and route-maps on CISCO routers representing ASes is carried out. In the end, an operationally deployable security policy is proposed which aims to protect BGP sessions and border routers from exploitation with little or no modification to the existing routing infrastructure.

Title *Page*

CHAPTER 3: RESEARCH METHODOLOGY

CHAPTER 5: CONCLUSION

APPENDIX A

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1 INTRODUCTION

1.0 INTRODUCTION

The internet is a global decentralized network of networks comprised of end systems that originate and or receive IP packets. These networks are identified by IP addresses as well as active forwarding elements called routers which forward IP packets through the network.

The internet started off as a US Department of Defense (DoD) network to connect scientists and university professors around the world (Peter, 2012). Currently, the internet serves as a global data communication platform linking millions of private, public, academic and business networks via international private leased circuits (IPLC) and optical fiber networks.

The internet has transformed the computer and the communications world like nothing before; providing opportunity for worldwide broadcasting; mechanisms for information dissemination and a medium for collaboration and interaction between individuals and their computers without regard for geographic location.

The internet consists of thousands of autonomous systems (AS) each owned and operated by a single institution (Sachdeva et al, 2011). Hence the internet can be said to be a conglomeration of autonomous systems that define the administrative authority and routing policies of different organizations. Autonomous systems are made up of routers that run interior gateway protocols such as Routing Information Protocol(RIP),Enhanced Interior Gateway Routing Protocol(EIGRP), Open Shortest Path First(OSPF), Intermediate system-Intermediate System(IS-IS) within their boundaries and interconnect via an Exterior Gateway Protocol(EGP).

The current internet de facto standard EGP is the Border Gateway Protocol Version 4(BGP-4) defined in RFC 1771 on 4 March 1995 by Rekhter et al (1995) and revised in RFC 4271 on 4 January 2006 by Rekhter et al (2006).

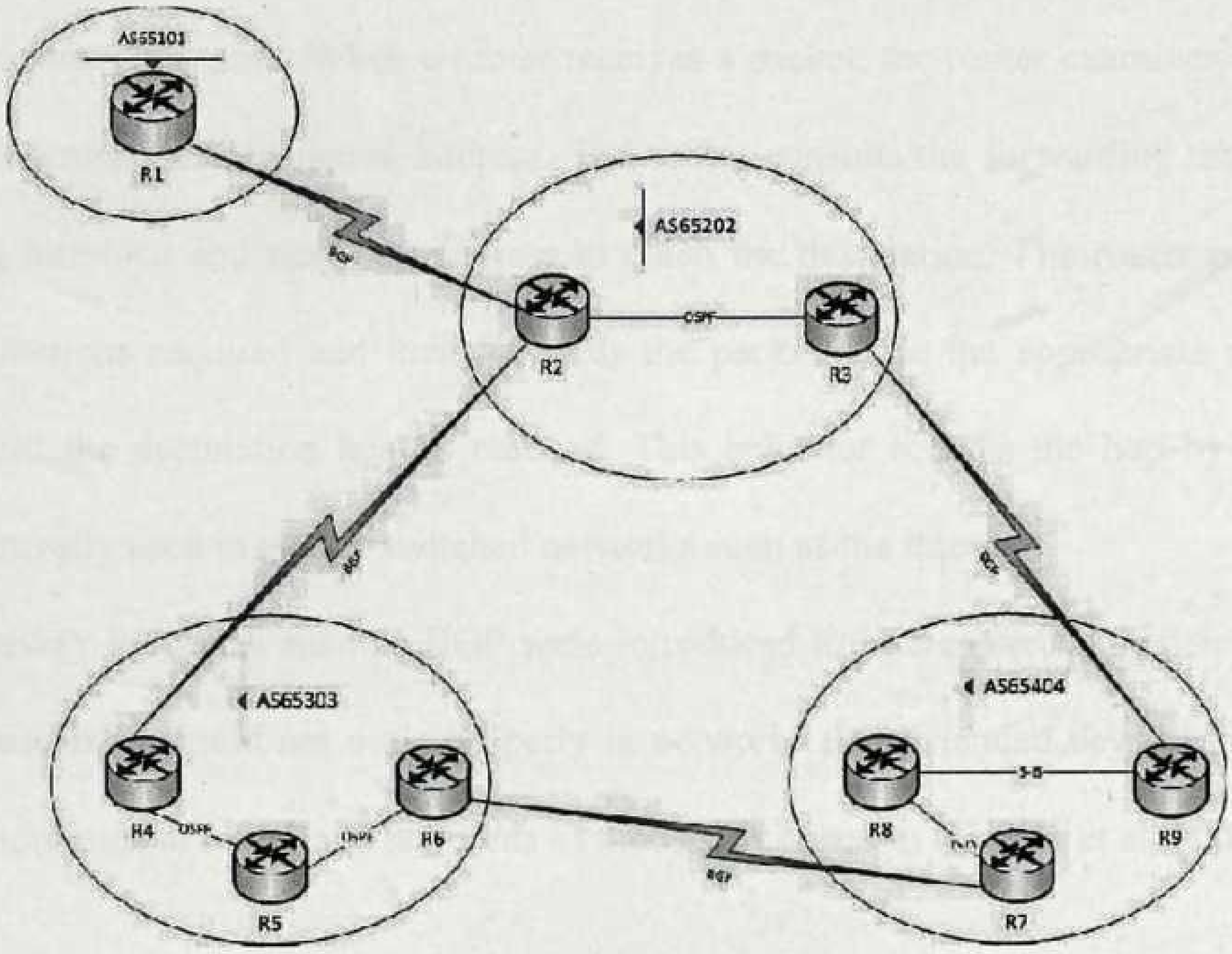

Figure 1.1 Autonomous systems running both EGP and IGP
Source: Sachdeva et al (2011)

Routers are devices that direct traffic between hosts. They build routing tables that contain collected information on all the best paths to all destinations that they know how to reach. Routers use routing protocols to both transmit and receive route information to and from other routers in the network. Routers use this information to populate routing tables that are associated with each particular routing protocol and consequently select the best path to each destination. Routers associate with that destination, the next hop device's attached data link layer address and the local outgoing interface to be used when forwarding packets to the destination. The next hop device in this case could be another router or even the destination host. The next hop device's forwarding information which is the data link layer address plus outgoing interface is placed in the router's forwarding table. When a router receives a packet, the router examines the packet's header to determine the destination address. The router consults the forwarding table to obtain the outgoing interface and next-hop address to reach the destination. The router performs any additional functions required and then forwards the packet on to the appropriate device. This continues until the destination host is reached. This behavior reflects the hop-by-hop routing paradigm generally used in packet switched networks such as the internet.

Exterior Gateway Protocols such as BGP were introduced to interconnect autonomous systems (ASes) because IGPs could not scale properly in networks that extended beyond the enterprise level, with thousands of nodes and hundreds of thousands of routes (Caesar et al, 2005).

The Border Gateway Protocol is the default inter domain routing protocol on the Internet which controls the packet forwarding behavior on a data plane and has significant impact on the well being of the global internet. BGP is a path vector protocol used to carry routing information between autonomous systems. The term path vector comes from the fact that BGP routing information carries a sequence of AS numbers that indicates the path of ASes that a network prefix has traversed. The path information associated with the prefix is used to enable loop prevention.

BGP uses TCP as its transport protocol on port 179 (Zhao et al, 2005). This ensures that all the transport reliability such as retransmission is taken care of by TCP and does not need to be implemented in BGP thereby eliminating the complexity associated with designing reliability into the protocol itself. Routers that run a BGP process are often referred to as BGP speakers. Two BGP speakers that form a TCP connection between one another for the purpose of exchanging routing information are referred to as neighbors or peers. Peer routers exchange open messages to determine the connection parameters. These messages are used to communicate values such as the BGP speakers version number while gracefully providing a mechanism to close a connection with a peer using a NOTIFICATION ERROR message.

Although the internet is globally accessible, users will have to connect through a local internet service provider (ISP) such as ZIPNET, K-NET in Ghana. Local ISPs in turn rely on regional IPLCs (International Private Leased Circuit) such the MAINONE CABLE COMPANY or SAT-3 to transport their traffic other national ISPs across continents.

National ISPs connect to each other through Network Access Points (NAP). Many ISPs these days however are connected to each other in what is known as ISP peering in order to achieve redundancy as well as to provide reliable Internet connectivity to clients.

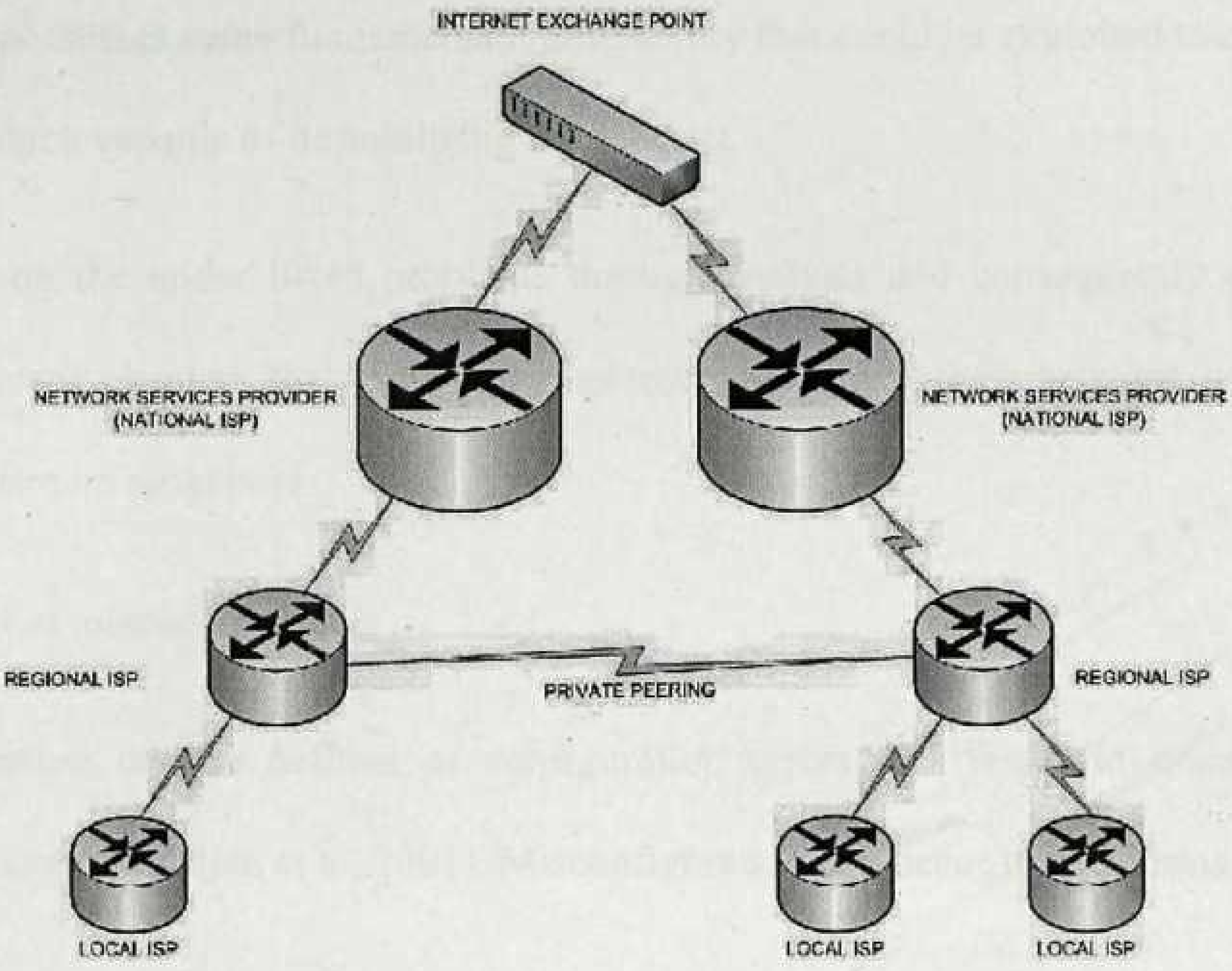

Figure 1.2 Levels of Internet Service Providers
Source: Caesar et al, 2005

1.1 STATEMENT OF THE PROBLEM

The internet started off as a research project; with the first few interconnected nodes being educational facilities. Due to the original intent for the development of the internet, not much attention was paid to security. The internet in present day has been leveraged to provide crucial services and applications. Notable among these is email, e-banking, e-commerce, telemedicine applications and transport control systems. These applications rely on the safety, scalability and reliability of the internet.

Exterior Gateway Protocols such as BGP logically bind the ASes that make up the internet together by providing a mechanism for BGP peers to exchange route information.

BGP unfortunately possesses some fundamental vulnerability that could be exploited to carry out different forms of attack capable of destabilizing the Internet.

This work focuses on the under listed problems through analysis and consequently proposes solutions in subsequent chapters that ISPs can implement to secure their sessions with their upstream and downstream neighbors.

1. The problem of misconfiguration

   Misconfiguration can be defined as configuration errors that result in unintentional leaking of routes (Mahajan et al, 2001). Misconfiguration can occur in two forms thus

   i. Origin misconfiguration

   ii. Export misconfiguration

2. Session hijacking

   Session hijacking is when an attacker places himself in between the source device and the destination device. This is also known as the man in the middle attack.

   BGP operates on trust. BGP speakers themselves inject bogus routing information either by masquerading as any other legitimate BGP speaker or by distributing unauthorized routing information as themselves. Historically, misconfigured and faulty routers have been responsible for widespread disruptions in the internet as the literature in subsequent chapters would demonstrate.

Although BGP has other vulnerabilities, the scope of this thesis will be limited to the previously mentioned in the problem statements section.

1.2 AIMS AND OBJECTIVES

This work sets out to find answers to a number of questions in the scope of the problem statement mentioned earlier and has at its core, the following specific goals.

1. Comparative analysis of the various proposals for BGP and identify why most of these suggestions have not been implemented in real life networks.
2. Assess the internet's routing infrastructure with focus on the design and operation of BGP.
3. Identify the sources and impact of misconfiguration in BGP updates and suggest possible counter measures.
4. Perform a simulation based demonstration of a deployable solution as a countermeasure to BGP session hijacking.
5. Finally, this work aims to demonstrate an advanced security policy that can be implemented to prevent BGP attacks with little or no changes in the existing BGP infrastructure.

1.3 SIGNIFICANCE OF THE STUDY

The internet is globally accessible and as a result, any major changes to its primary routing architecture could initiate a series of configuration changes in end systems and a possible disruption of the global internet.

> The practical security policy advanced in this work would provide service providers an alternative that they could adopt to mitigate the effects of BGP's vulnerabilities and also prevent their networks from being used to launch attacks against other BGP peers and ASes without requiring any major changes in the routing infrastructure.

BGP is essential in today's Internet through it's inter domain routing functionalities. A security policy that helps strengthen BGP security between ASes could help improve the overall security and availability of the global internet.

1.4 DESCRIPTION OF CHAPTERS

This thesis is divided into five main chapters. This subsection provides an overview of each chapter as presented in the thesis.

CHAPTER 1: INTRODUCTION

This chapter provides a brief background history. It goes on to talk about internet routing and also introduces BGP. The problems to be addressed in this work are also stated as well as the aims and objectives for the study. The significance of the study is also included in this chapter.

CHAPTER 2: LITERATURE REVIEW

The security measures proposed in literature is analyzed and the relevance of the intended research established by pointing out the loop holes in the existing tools and policies proposed for BGP security by other researchers.

CHAPTER 3: METHODOLOGY

This chapter focuses on the methods to be used in conducting this research. It sets out the type of research to be conducted as well the tools for implementing and testing the proposed security solutions. Simulations for the session hijacking problem are carried out here.

CHAPTER 4: BGP MISCONFIGURATIONS

This chapter presents the findings on the study of BGP misconfiguration. It identifies the causes and effects of the various forms of misconfigurations and proposes mitigation measures that service providers could implement to arrest the situation

CHAPTER 5: CONCLUSIONS, RECOMMENDATIONS AND FUTURE WORK

This chapter concludes the work based upon the findings of the research and puts forward a couple of recommendations. Suggestions for future research in the area of BGP security are also included in this section.

CHAPTER TWO LITERATURE REVIEW

2.0 OVERVIEW

The internet is a decentralized collection of interconnected computer networks. These networks are composed of end hosts and active forwarding elements whose role is to pass IP packets through the network.

The routing system distributes the relative location of addresses to each routing element for consistent and optimal routing decisions in order to pass traffic from the source to destination using Routing protocols.

The Border Gateway Protocol (BGP) (Rekhter et al, 2006), RFC 4271 has provided inter-domain routing services for the Internet's disparate component networks since the late 1980s. The central role of routing that BGP performs makes it one of the most critical protocols that provide security and stability to the Internet. (Office of the US President, 2004)

The inter-domain routing infrastructure is still susceptible to attack. (Office of the President, US 2004) As a result of trust relationship between peers, BGP has no build in peer and update authentication mechanisms. This trust model could be exploited by attackers to destabilize the global internet. Rogue ASes could advertise prefixes they don't own deliberately to launch attacks or redirect traffic. (Murphy, 2006), (Babier et al, 2006) (Ballani et al, 2007) (Zetter, 2008)

This could lead to the failure of key internet services such as email and render certain ASes or websites artificially unavailable. The effects could be anything from minor degradation of application performance to major disruption (Underwood, 2006), (Brown, 2008). Currently research on BGP is focused on two major issues thus scalability and security to ensure integrity of BGP updates. The chapter focuses on the security of BGP through a comparative analysis of the various tools and approaches available for securing BGP – pointing out their strengths and weaknesses; while bringing to light the relevance of the intended research.

2.1 SECURING BGP

The vulnerabilities of BGP are borne out of four fundamental weaknesses in BGP and the inter-domain routing infrastructure according to Murphy (2006). These are

i. Absence of mechanisms to protect the integrity, freshness and source authenticity of BGP messages.
ii. Absence of mechanisms to verify the originality of an address prefix and the ASes that originate the prefix.
iii. Absence of mechanisms to verify the originality of attributes of a BGP UPDATE message.
iv. Absence of mechanisms to verify that the local cache RIB information is consistent with the current state of the forwarding table.

The above mentioned factors represent the specific facets of BGP research with respect to security. A lot of the suggestions reviewed in this chapter aim at addressing a number of these specific observations.

2.1.1 THE SECURITY TOOLSET

A number of mechanisms for securing BGP have been developed which begin at the session level and also includes the tools that are used to protect the TCP session at both the sending and receiving end.

According to Gill et al (2004) the TTL security mechanism is one such proposal that could substantially limit the effective radius of potential attack on the session.

There are two tools to protect the BGP TCP session from external disruption that rely on the use of a cryptographic function.

These are the use of IPSEC at the IP proposed by Kent et al (1998) and the TCP MD5 signature option at the TCP session level proposed by Rivest (1992) and revised by Hefferman (1998).

The MD5 signature option has some potential weaknesses when compared with IPSEC based on the assessment of Murphy (2006) however the MD5 signature option is preferable to no form of TCP protection at all. The choice between IPSEC and MD5 is made by considering their key relative capabilities. No standard key rollover mechanism exists in MD5 as asserted by Behringer (2007) alongside the cryptographic processing load it comes with; whereas the load of IPSEC processing is significantly higher than MD5 processing.

The cryptographic validation requirement of these two mechanisms provides room for a potential denial of service threat where a BGP speaker could be flooded with invalid messages each of which must be cryptographically processed before being detected as invalid and discarded (Christian et al, 2008). In addressing the message integrity limitation, an approach is suggested by Schneier (1995) which aims to provide transparent session level protection through the use of digital signatures. By this mechanism, a set of credentials is assigned that allows peers to verify the correctness of the information carried as the message payload in BGP.

The reason for the using of digital signatures instead of an integrity check which uses some form of shared secret key is due to the fact that the number and identities of all external recipients of the information is not known in advance (Murphy, 2001).

Apart from being able to determine whether or not a message had been altered en route to the destination, a mechanism to actually verify the authenticity of the original information is necessary.

This meant that the digital signatures used had to be verified thus using some form of mechanism that authenticates the public key associated with an address prefix or an AS number (Cooper et al, 2008)

2.1.2 APPROACHES TO SECURING BGP

A very significant contribution to this area is the secure BGP (SBGP) proposal by Kent et al (2000). This happens to be one of the most complete contributions in this direction despite the fact that the assumptions relating to the processing capabilities of the routing equipment needed to run the protocol far exceeds what is available in real life.

SECURE BORDER GATEWAY PROTOCOL (SBGP)

The sBGP protocol places digital signatures over the address and AS path information contained in routing advertisements and defines an associated PKI for validation of these signatures. sBGP defines the correct operation of a BGP speaker in terms of constraint placed on individual protocol messages, including ensuring that all protocol UPDATE messages have not been allowed in transit between the BGP peers and that the UPDATE messages were sent by the indicated peer. The basic security framework proposed in sBGP is that of digital signatures thus x.509 certificates and PKI's. This enables BGP speakers to identify and authorize other BGP speakers as well as AS administrators and address prefix owners. The verification framework for sBGP requires a PKI for address allocation, where every address assignment is reflected in an issued certificate (Seo et al, 2000). This PKI provides a means of verification of a "right-of-use" of an address. In addition, sBGP proposes the use of IPSEC to secure the inter-router communication paths as well as the use of attestations. An address attestation is produced by an address holder, and authorizes a nominated AS to advertise itself as the origin AS for a particular address prefix.

There are a number of significant issues that have been identified with sBGP including the computation burden for signature generation and validation as well as the increased load in BGP session restart. There is also the issue of piecemeal deployment and the completeness of route attestations (Zhao et al, 2005)

SECURE ORIGIN BORDER GATEWAY PROTOCOL (soBGP)

A refinement to the sBGP approach is secure origin BGP (SoBGP) proposed by White (2003) in an effort to find a middle ground between the additional security processing overhead and the capabilities of deployed routing systems and security infrastructure. Here, the requirements for AS path verification are relaxed and the nature of the related Public Key Infrastructure is altered to remove the requirement for a strict hierarchical address PKI that precisely reflects the address distribution framework.

soBGP uses the key concept of an Entity Cost to bind an AS to a public key. SoBGP avoids the use of hierarchical PKI that mirrors the AS number distribution soBGP uses the concept of an AuthCert to bind an address prefix to an originating AS. This AuthCert is not signed by the address holder, but by a private key that is bound to an AS via an EntityCert. The explicit avoidance of reliance on the established AS and address distribution framework and any form of associated PKI as the derivation of a trust hierarchy may have been a pragmatic consideration in the design of this approach, but it leaves open the issue of how to establish trust anchors for validation of these signed objects. This is a rather significant deficiency in the validation framework for soBGP

The overall approach proposed in soBGP represents a different set of design trade-offs to sBGP, where the amount of validated material is a BGP UPDATE message is reduced. This can reduce the processing overhead for validation of UPDATE messages. In soBGP each local BGP speaker assembles a validated inter-AS topology map as it collects ASPolicyCerts, and each AS path in UPDATE messages is then checked to see if the AS sequence matches a feasible inter-AS path in this map. The avoidance of a hierarchical PKI for the validation of AuthCerts and EntityCerts could be considered a weakness in this approach, as the derivation of authority to speak on addresses is very unclear in this model.

PRETTY SECURE BGP (psBGP)

Another refinement of the SBGP model is pretty secure BGP (psBGP) proposed by Oorschot et al (2007). This approach represents a similar effort aimed at achieving a compromise between security and deployed capability through the introduction of a trust rating for assertions based on assessment of confidence in corroborating material.

psBGP puts forward the proposition that the proposals relating to the authentication of the use of an address in a routing context must either entirely rely on the use of signed attestation that need to be validated in the context of PKI, or rely on the authenticity of information contained in Internet Routing Registries . The weakness of routing registries is that the commonly used controls in the registry are insufficient to validate the accuracy or the current authenticity of the information that is represented as being contained in a route registry object.

psBGP allows for partial path signature to exist, mapping the validation outcome to a confidence level rather than a more basic BGP model of accepting an AS path only if the AS path in the BGP UPDATE is completely verifiable. The essential approach of psBGP is the use of a reputation scheme in place of a hierarchical address PKI, but the value of this contribution is based on accepting the underlying premise that a hierarchical PKI for addresses is feasible. psBGP appears to be needlessly complex and bears much of the characteristics of making a particular solution for the problem, rather than attempting to craft a solution within the bounds of the problem space.

INTERDOMAIN ROUTE VALIDATION (IRV)

Another technique, inter-route validation (IRV) proposed by Goodell et al (2003), attacks the problem from a different angle by extending the existing model of Internet Route Registries into per-AS route registries. It attempts to replace the configuration of the BGP protocol with security credentials, in a query based credential retrieval system. The approach assesses the security function as an incremental overlay on the existing routing infrastructure.

This approach is midway between the strict AS path test of sBGP that validates that the UPDATE message was passed along the AS sequence described in the AS Path and the soBGP AS Path feasibility that validates that there is a set of AS peer connections that correspond to the AS sequence.

Here the validation test is that each AS in the sequence is currently advertising this prefix to the next AS in sequence.

This IRV architecture has a number of issues that are not completely specified, including IRV discovery, IRV query redirection, authentication of queries and responses, selective responses, transparent layer protection and imposed overheads.

It is unclear how an IRV response is to be validated, and how the relying party can verify that the received response originated from the IRV server of the AS in question, that the response has not been altered in any way, and that the response represents the actual held state in the queries AS. A similar concern lies in the estimation of additional overhead associated with performing a query to each AS in the AS Path for every received BGP update. It is also unspecified whether the query and response is a pre-condition for acceptance of a route would appear to offer a route robust form of security, it is also the case that IRV would be unreachable until the route is accepted

2.2 SECURING THE INTEGRITY OF BGP DATA

Another proposal that addresses routing security is the 1998 study on Byzantine Robustness by Perlman (1998). When one or more routing entities fail or exhibit malicious behavior, it is expected that all other properly working entities would attain mutually consistent decision in light of the validity of each message within a defined state of time. The study was in the area of link-state protocol design, and the work described a protocol that satisfied the properties for Byzantine Robustness. This approach typically did not match the inter-domain routing environment as a result of its link state focus however the concept of validation of routing information permeates all BGP Security Architectures.

Smith and Garcia-Luna-Aceves (1996) approach the session security requirements by proposing modifications to the BGP protocol with the introduction of message encryption at the BGP level. This approach involves the exchange of session keys when the BGP session is established as well as the addition of a message sequence number that protects against replay attacks and message removal. There is also the introduction of a new attribute called the predecessor path attribute showing the AS prior to the destination AS as well as the digital signing of all fixed fields in the UPDATE message. The predecessor attribute is meant to validate the AS path attribute. The proposed changes to the BGP protocol required comprehensive adoption and deployment in order to be effective.

Another proposal for securing BGP proposed by (Bales et al, 1998) focused on allowing the BGP data flow to access credential information that allows a BGP speaker to confirm the authenticity of origination information in BGP update messages by validating the binding of address prefixes to originating ASes. This work proposed the use of DNS as the distribution mechanism allowing a BGP speaker to perform a DNS query to authenticate the origination provided in a BGP UPDATE message. This approach was never deployed due to limitations in the DNS structure itself. For example an entire address block could be sub-allocated to a DNS sub-level causing a prefix to refer to the DNS entry that did the "allocation in full" instead of the legitimate owner of the prefix

2.3 CRYPTOGRAPHIC APPROACHES TO BGP SECURITY

Hu (2003) asserted that symmetric cryptographic techniques such as message authentication codes (MACs) or cryptographic hash functions are 3 to 4 times faster than asymmetric cryptographic functions when applied to digital signatures (Hu et al, 2003). The high cost associated with the use of asymmetric cryptographic functions for the authentication of AS Path information is the major limitation preventing the adoption and deployment of these techniques.

The Secure Path Vector routing (SPV) for securing BGP is a proposal that explores the possibility of using symmetric cryptography in securing the AS Path in BGP update messages (Hu et al, 2004) as an extension to hash chains functions. SPV applies tree authenticated hash values to AS Path validation as an alternative to the nested digital signature proposed as the AS path validation mechanism for SBGP.

This proposal belongs in the category of proposals that call for changes to the operation of the BGP protocol. The significant change is that all routes must be re-advertised to pass within a fixed time intervals. This affects performance adversely making the protocol unattractive for commercial deployment. Raghavan et al (2007) also argue that SPV couldn't avert route forgery, eavesdropping or collusion attacks to re-enforce the earlier point.

There is no clear cut solution to the problem of routing security that attains a balance between security and acceptable deployment overhead (Chan et al, 2006). Current research on BGP security focuses on the integrity, authenticity, and verifiability of routing information (Butler et al, 2010). A more stable routing system capable of providing stable routing state is also capable of verifying routing information updates.

2.4 CONCLUSIONS

It is against this backdrop that this research is conducted in an attempt to put forward a security policy based on best practices that service providers can employ to counter act the specific problems identified in Chapter 1.

We believe the solution to BGP session hijacking and misconfiguration does not necessarily rest with protocol extensions or modifications but rather the implementation of policy based routing such as the use of filters, route-maps and policy maps as well as a number of already known best practices in the service provider industry. This approach will require no additional hardware but what exist already and should guarantee a level of BGP session security at service provider level if implementation is efficiently done.

Unlike the earlier suggested approaches to securing BGP, this approach focuses on the optimization of BGP configuration using Defensive Routing Policies from Cisco IOS to create a more robust BGP session better equipped to forestall session hijacking attacks while taking into consideration the processing capabilities of the existing routing infrastructure or equipment in other to prevent undue overhead associated with the implementation of other proposals for securing BGP. The approach proposed will also include workable everyday practices that service providers can follow to beef up the overall security of their core routing equipment.

CHAPTER 3 RESEARCH METHODOLOGY

3.0 INTRODUCTION

This chapter describes the approaches adopted for the intended study. It provides some background information on the design and operation of BGP together with internet routing while focusing on the general guidelines for problem solving and experimentation. The chapter provides specific descriptions of methods; techniques and tools employed for the simulations as well the sources and methods for data collection and analysis.

This work focuses on counter measures against BGP session hijacking and misconfiguration. The methodologies adopted in this research are directly influenced by the specific problems the thesis attempts to address.

3.1 DESIGN AND OPERATION OF BGP

The Border Gateway Protocol (BGP), defined in RFC1771 (Rekhter et al, 1995) and refined in RFC4271 (Rekhter et al, 2006) allows the creation of loop free inter domain routing between autonomous systems. A set of routers under a single technical administration constitutes an autonomous system (Caesar et al, 2005).

Routers in AS may use multiple interior gateway protocols to exchange routing information inside the AS and an exterior gateway protocol to route packets outside the AS as discussed in earlier chapters.

3.1.1 OPERATION OF BGP

BGP employs TCP as its transport protocol on port 179 (Zhao et al, 2005). Two BGP speaking routers form a TCP connection between one another. These two may be referred to as PEER ROUTERS. Peer routers exchange messages to open and confirm the connection parameters.

BGP routers exchange network reach-ability information which indicates full paths that a router should take in order to reach the destination network. This information is used in the construction of AS graphs that provide indications as to where routing policies can be applied in order to enforce restrictions on the routing behavior.

3.1.2 BGP and TCP

BGP uses TCP as a reliable transport protocol to support the protocol's transactions across a BGP peer session (Huston et al, 2011). Essentially, BGP assumes that a link is always established for the forwarding of IP packets at the link level. TCP ensures reliable message delivery as well as flow control between the BGP. The TCP stream is divided into messages using BGP-defined markers, where each message is between 9 and 4096 octets in length. The use of a reliable transport platform implies that BGP need not explicitly confirm receipt of a protocol message. This removes much of the protocol overhead seen in other routing protocols that sit directly on top of a media level connection. There are no message identifiers, no message number initiation protocol, no explicit acknowledgement of messages or any provision to manage lost, re-ordered or duplicated messages. These functions are managed by TCP and are therefore unnecessary for BGP to also support. The use of a reliable transport protocol also obviates the need for BGP to periodically refresh the routing state by re-flooding the entire routing information set between BGP speakers. After the initial exchange of routing information, pair of BGP routers exchanges only incremental changes to routing information.

3.1.3 BGP PEERS

Any two routers that have formed a TCP connection in order to exchange BGP routing information are called PEERS (Cisco Press, 2008). They could also be referred to as neighbors.

BGP peers initially exchange their full BGP routing tables. Incremental updates are sent as the routing table changes. BGP keeps a version number of the BGP table and it should be the same for all of its peers. The version numbers changes whenever BGP updates the table due to some routing information changes. Keep-alive packets are sent to ensure that the connection is alive between the BGP peers and notification packets are sent in response to errors or special conditions.

3.1.4 EBGP and IBGP

An autonomous system that has multiple BGP speakers could be used as a service for other AS's

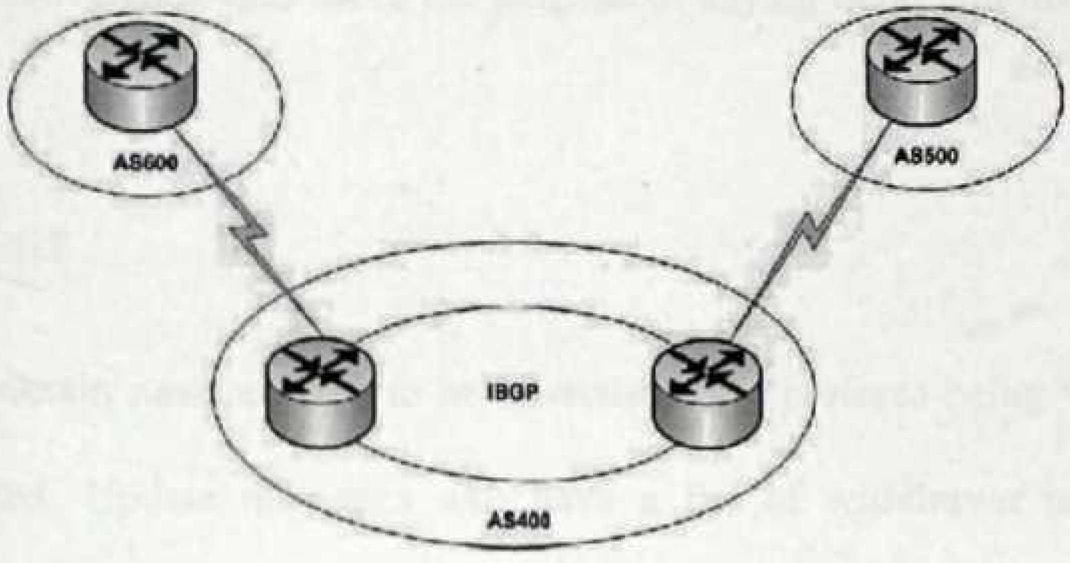

Figure 3.1 AS 400 provides transit services to AS 600 and AS500.
Source: Cisco Press (March, 2008)

It is important to ensure reachability for networks within an AS before sending the information to other external AS's. EGBP refers to BGP running between two routers belonging to two different AS's while IBGP refers to BGP running between routers in the same AS(Cisco Press, 2008).

3.1.5 BGP MESSAGES

The BGP routing protocol exchanges different messages for co-coordinating between participating routers in BGP domain.

OPEN MESSAGES

Open message contains BGP version number, AS number and hold down time that need to be negotiated with neighbor before peer relationship.

KEEP ALIVE MESSAGE

Once peer relationship is established, the peer routers will keep on sending keep-alive messages to each other. Keep-alive messages serve the purpose of saying hello and informing peers about each other's status.

UPDATE MESSAGE

Update messages contain new prefixes to be advertised and prefixes being withdrawn that was previously advertised. Update messages will have a list of withdrawn prefixes or/and new prefixes with their path attributes.

NOTIFICATION MESSAGE

Notification message is sent whenever some error or mismatch is detected, notification message is sent to peer and connection is closed. The very important point here to note is that whenever a peer relationship is broken between two peers, all routes learned through said peer are purged but from the routing database.

3.1.6 BGP STATES

BGP process running on routers moves through different states during its execution.

IDLE STATE

Whenever router is not configured for BGP, BGP process on this router is in idle state waiting for statement which is the manual peer configuration of BGP. Once BGP process is started on router, it initiates TCP resources, starts listening and moves to CONNECT STATE.

CONNECT STATE

In connect state, BGP process waits for establishment of TCP connection after which OPEN message will be sent for peer negotiation and BGP process moves to open SENT state. If TCP connection could not be established, the router moves to ACTIVE state and waits for connect retry timer to expire before attempting the TCP connection again.

ACTIVE STATE

In active state, the router is actively trying to establish TCP session with peer then after OPEN MESSAGE is sent and router transits to OPEN SENT state.

OPENSENT STATE

Open sent state means the router has sent an OPEN message containing its BGP parameters and waiting for the corresponding OPEN message from peer to negotiate and establish BGP peer relationship. The connection type is determined here, i.e. EBGP or IBGP, hold-down times negotiated and keep alive messages exchanged and state transits to OPEN CONFIRMED. If the corresponding OPEN message received has some mismatched or if the NOTIFICATION message is received, state transits to idle.

OPEN CONFIRM STATE

Open confirm state means BGP peering relationship has been established between the peers and they have started exchanging keep-alive messages. In this state if keep alive is received, the BGP process on that router moves to ESTABLISHED STATE and hold down timer is started. If NOTIFICATION message or TCP reset is received, BGP router moves to IDLE STATE.

ESTABLISHED STATE

The established state is the final state where routing information is exchanged among peers. If notification message, TCP reset is received or keep alive is not received in due time and hold down timer express, the connection is dropped and the router moves to IDLE state.

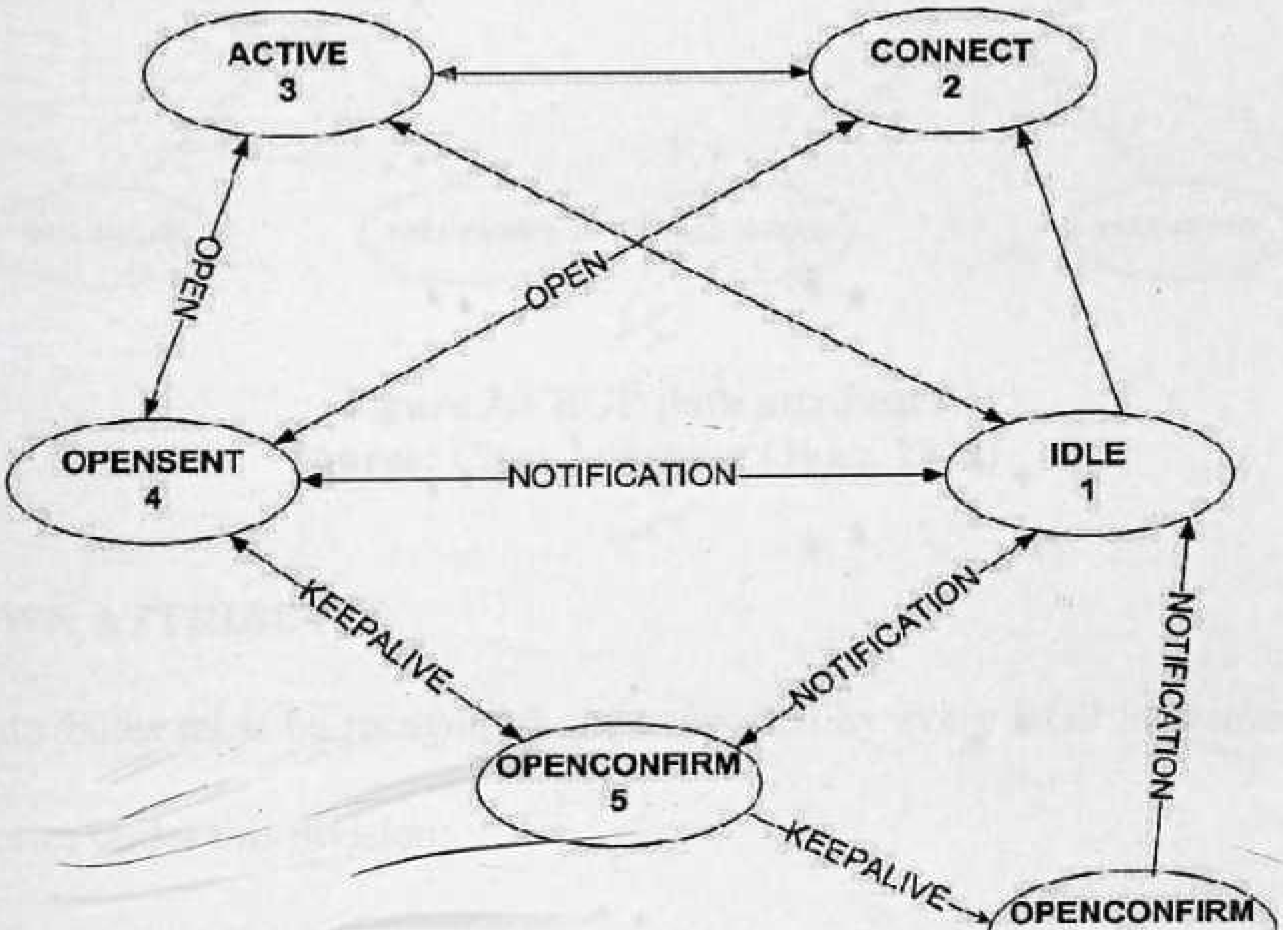

Figure 3.2 BGP finite state machines
Source: Cisco Press (March, 2008)

3.1.7 BGP PATH ATTRIBUTES

Path attributes are characteristics associated with BGP routes being advertised. These path attributes make BGP metric; BGP doesn't use simple metric like cost in OSPF. However BGP path selection involves a complex algorithm that makes decisions on the basis of these path attributes. There are different categories of path attributes as illustrated below.

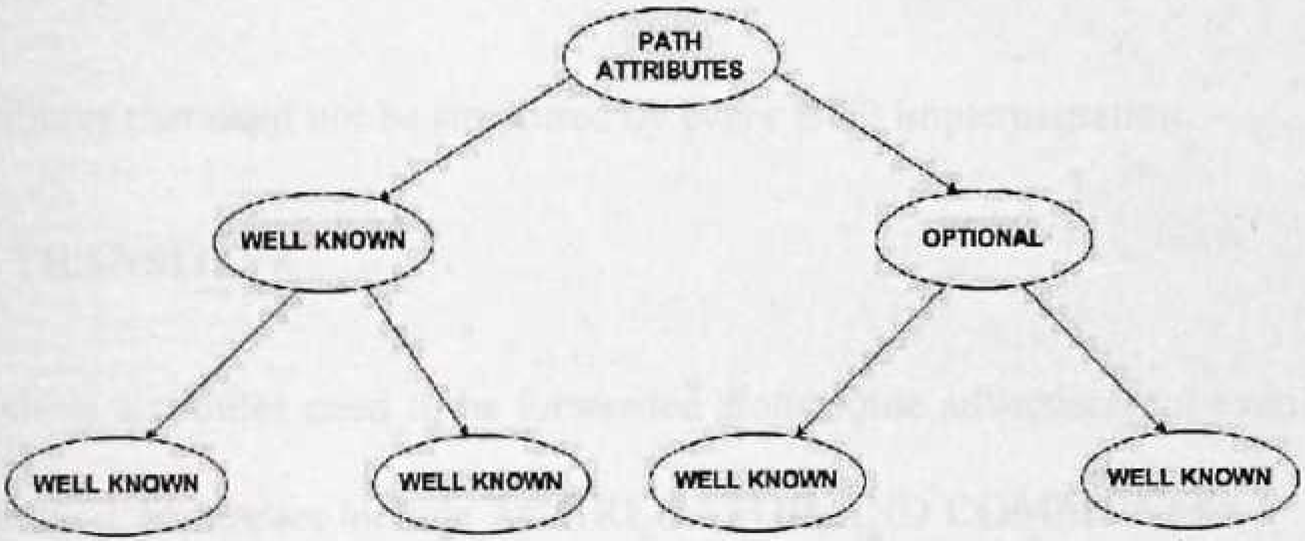

Figure 3.3 BGP path attributes
Source: Cisco Learning (June, 2012)

WELL KNOWN ATTRIBUTES

Well known attributes must be recognized and supported by every BGP implementation. There are two categories under this division:

i. Well known mandatory

ii. Well known discretionary

WELL KNOWN MANDATORY

Well known attributes are supported by every BGP router and well known mandatory attributes are included in every BGP route advertisement including ORIGIN, AS-PATH and NEXT-HOP.

WELL-KNOWN DISCRETIONARY

Well known discretionary attributes must be supported and understood by every BGP speaking router but it is not necessary to include them in every advertisement.

OPTIONAL ATTRIBUTES

These are attributes that need not be supported by every BGP implementation.

OPTIONAL TRANSITIVE

Optional transitive attributes need to be forwarded along route advertisement even if that router doesn't recognize it. Examples include **AGGREGATOR** AND **COMMUNITY**

OPTIONAL NON TRANSITIVE

These are optional attributes that can be deleted by a router if the router doesn't recognize them. An example is the Multi Exit Discriminator (MED).

AS-PATH ATTRIBUTE

The AS-PATH attribute is a well-known mandatory attribute associated with every BGP route and contains the list of autonomous system that must be transited to reach the prefix advertised by this route advertisement with originating AS listed at last. BGP is a path vector routing protocol, as by default the route having the shortest path is selected if no other setting is manipulated. The AS-path attribute is used for loop prevention in BGP. In that when a router receives NLRI (network layer reachability information) it will simply ignore this path considering it as possible loop.

THE ORIGIN ATTRIBUTE

The ORIGIN attribute defines how the associated route is originated at first place. ORIGIN attribute may have different values of origin code.

IGP: NLRI learned through configuration of network command will have origin code of IGP.

EGP: NLRI learned through EBGP are marked with origin of EGP.

INCOMPLETE: A route learned through redistribution from some other routing protocol is assigned with ORIGIN attribute value of incomplete.

THE NEXT HOP ATTRIBUTE.

The NEXT-HOP attribute associated with NLRI refers to next-hop IP address where the traffic should be directed to reach destination advertised by the NLRI. In IBGP, the next-hop is not always the peer router's exit interface IP address.

When advertising to different autonomous systems, the next-hop is the IP address of routers outgoing interface when advertising to the same autonomous system (IBGP) peer and the prefix is internal, next hop is outgoing interface IP address of the router that originated this advertisement.

When advertising a prefix that is learned from another AS to an internal peer, the NEXT-HOP is the IP address of the outgoing interface of the external AS peer from which NLRI was learned first.

THE LOCAL_PREF ATTRIBUTE

The local preference attribute is used inside autonomous systems to set and communicate the local preference for certain routes. If a router has two routes to the same destination, a route having higher local preference is preferred over the one that has lower local preference value.

The boundary router of autonomous systems assigns the local preference value and then it is communicated inside the autonomous system, so that all routers within the autonomous system can take the preferred path.

As shown in figure 3.4, RTR-0 receives two routes to network 192.168.1.0/30 one from RTR-1 and one from RTR-2. Routes learned through R1 have been assigned local preference of 250. When RTR-0 has to send traffic to 192.168.1.0/30, it will prefer the link B as it has higher local preference.

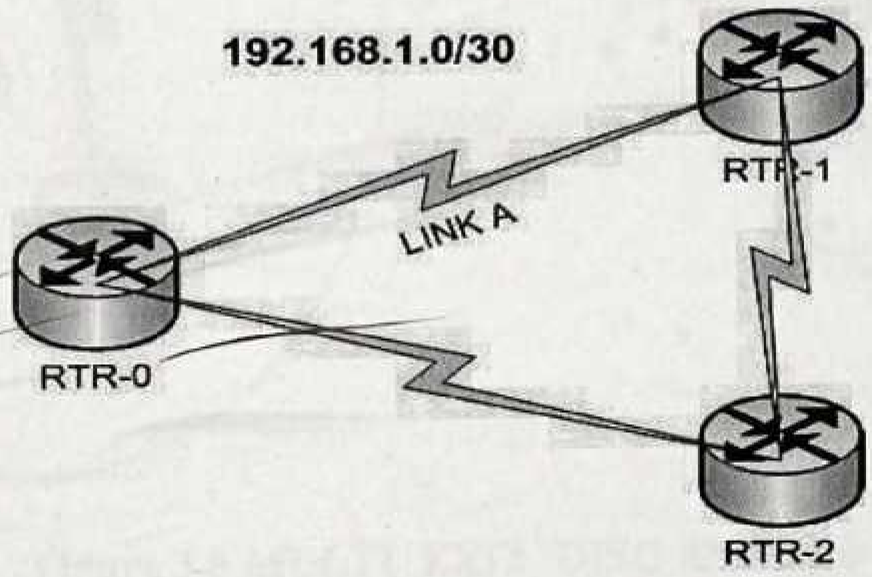

Figure 3.4 LOCAL-PREF attribute

Source: Rexford et al (2005)

MULTI_EXIT_DISC ATTRIBUTE

BGP has total control on traffic leaving the network and can enforce carry traffic policy using path attributes but have very little control over inbound traffic.

MED is the attribute through which ISP's can suggest a possibly better way towards it from the outside world, it is not binding on the receiving external peer to obey suggested MED. Lower value MED is preferred.

As shown in figure 3.5, an organization running BGP under AS number 100 and has internal network 192.168.1.0.Border router of the AS will advertise this same network to ISP with MED value 200 on link B. now the ISP has two paths towards internal network and it may choose the one with a lower MED thus through link A.

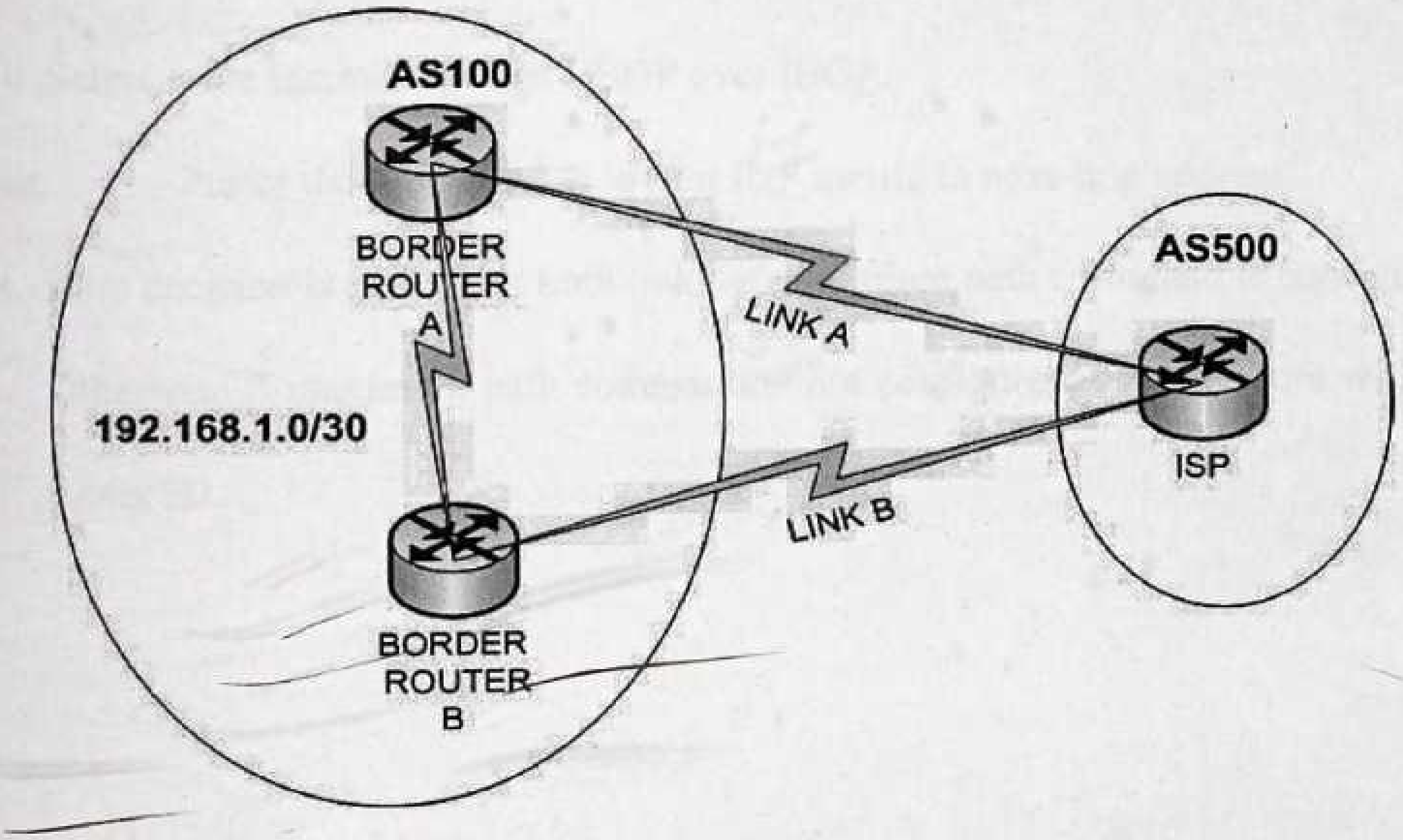

Figure 3.5 MULTI_EXIT_DISC Attribute
Source: Cisco Press (March, 2008)

3.1.8 BGP ROUTE SELECTION ALGORITHM

The BGP route selection algorithm as contained in Cisco press release (2008) is presented here. The first step is to check the availability of next-hop IP address, if next-hop is not reachable, the route is not considered. The algorithm is as follows:

i. The route having the highest weight is preferred.

ii. If two routes have equal value of weight attribute, local-preference attribute is checked and one having a higher local-preference value is preferred.

iii. If local-preference is equal, check if one of the routes is learned through IGP then it is preferred.

iv. Route having shortest AS-path is preferred, if a decision is not made in previous steps.

v. Check origin of route, IGP is preferred over EGP and INCOMPLETE is least preferred.

vi. If multiple route line up to that port, choose the route with the lowest MED value.

vii. Select route learned through EBGP over IBGP.

viii. Prefer the route having lowest IGP metric to next-hop address

ix. If no decision is made, use both links, if maximum path command is configured.

x. Otherwise if maximum path command is not configured, use the route with lowest BGP router ID

3.2 RESEARCH DESIGN

This work attempts to address two specific security vulnerabilities of the Border Gateway Protocol namely:

a) The problem of misconfiguration

b) The problem of session hijacking

To address these two issues however different approaches, tools, data sets, case studies and scenarios are required. This makes the intended study one which employs a concatenation of variables of various methodologies to arrive at one holistic security policy that service provides can use to further secure their BGP sessions.

A stepwise approach that involves experimentation by way of simulations for the session hijacking problem as well as the use of secondary data for the analysis of BGP misconfiguration.

3.3 BGP MISCONFIGURATION

It is a well known fact that simple, accidental BGP configuration errors can disrupt internet connectivity, yet little is known about the frequency of misconfiguration or its causes (Mahajan et al, 2001)

The data for the analysis is secondary data for the month of April 2012 obtained from Route-views which is a public looking glass. The data is generated by analyzing the routing table advertisements from 23 vantage points across the internet backbone while looking out for incidents of misconfiguration. For the incidents identified, the administrators at Route-Views tried to contact the ISPs involved through e-mail to verify whether it was a misconfiguration and to learn the cause of the incident. The analysis and perspectives put across on the subject matter (misconfiguration) is based largely on the data available and as such all assumptions made by the data source by extension may be carried through to this thesis as well. The assumptions made were used as a guiding framework to identify and distinguish BGP misconfiguration from other types of errors on the global routing tables.

Assumption 1: BGP policy changes typically operate on human time-scales, while changes due to misconfiguration and failures typically last for a much shorter time.

Assumption 2: A new route is either a new prefix or a new origin for an existing prefix.

It was observed that most new routes either last less than a day or last much longer; forty-five percent (45%) of the changes last less than a day, while thirty percent (30%) of them lasted more than 7 days.

Misconfiguration incidents are then identified by focusing on the short lived changes that last less than a day, the nature of the change depends on the type of misconfiguration.

Some of the short-lived changes identified as potential misconfiguration could also have been caused by legitimate events therefore the email survey of operators involved in each incident were conducted to confirm misconfiguration.

LIMITATIONS: These are limitations that militate against the data gathering methodology which must be taken into account as it affects the overall accuracy or otherwise of the research findings. The study underestimates the extent of misconfiguration in several ways.

i. Only misconfigurations that last less than a day are considered. Hence errors that persist for a longer time are missed.

ii. Not all kinds of misconfiguration events are considered.

iii. Only misconfiguration events that reach the Route-Views server despite filtering and best path selection at each hop are observed.

For these seasons, the result should be considered as a lower bound on the extent of BGP misconfiguration.

3.4 SESSION HIJACKING

Session hijacking involves intrusion into an ongoing BGP session by masquerading as a legitimate peer in a BGP session (Behringer, 2007). The difference as compared to a TCP reset attack is that session hijacking attack may be designed to achieve more than simply bringing down a session between BGP peers. The objective may be to change routes used by a peer or black holing (IETF, 2006), (RFC 4272)

In EBGP, neighbor routers add all routes they receive from their downstream neighbors into their routing table and then advertise those routes to the next hop. BGP session hijacking could occur when ISPs do not filter advertisements. An attacker could then hijack such an ISP and use their routers to advertise any prefix they want leading to a diversion of traffic or a blackhole as the simulations would demonstrate.

Session hijacking attacks could result in serious outages culminating in a complete loss of connectivity. Specific instances include the 2008 incident where at least eighty US universities had their traffic diverted to block access to their site from inside the country but accidentally black holed the route in the global BGP table (Docstoc, 2011).

In studying the session hijacking problem, case studies were developed that feature a number of routers representing ISPs and other autonomous systems connected together through EBGP peering. To successfully carry out the attack, a rogue AS (autonomous System) was included in the case study to advertise legitimate routes to its downstream neighbors creating a masquerade effect that causes downstream neighbors to forward traffic towards the rogue AS with the fake identity.

Routes forwarded to this AS by its downstream neighbors are not forwarded to the next hop address (router) consequently creating a Black Hole.

3.4.1 SESSION HIJACKING EXPERIMENT.

As a result of the scalability and flexibility offered by simulators as well as the number of routers required in the experiment, the simulation approach was thought to be more feasible as opposed to using real life.

A number of Network simulators were evaluated for the experiment. These include GNS 3, NS-2 and OPNET.

GNS-3 (Graphical Network Simulator-3)

GNS-3 is a graphical simulator that allows users to design complex network topologies. It allows for the creation of simulations and the configuration of devices such as workstations and powerful simulations, GNS 3 is strongly linked with.

- **i.** Dynamic, a CISCO IOS emulator
- **ii.** Dynagen, a text-based front end for Dynamics
- **iii.** Qeme, a generic and open source machine emulator.
- **iv.** Virtual Box which is a free and powerful virtualization software

GNS-3 can also be used to experiment features of the CISCO IOS, JUNIPER JUNOS or to cheek configurations that need to be deployed later on real routes. Ultimately GNS-3 is an open-source, free program that may be used on multiple operating systems including Windows, LINUX and MAC OS X.

NS-2 (Network Simulator 2)

NS-2 is a discrete event simulator targeted at networking research NS-2 supports the simulation of TCP, routing and multicast protocols over wired and wireless networks.

Despite the confidence deposed in NS-2 by researchers, NS-2 still remains a work in progress. NS-2 therefore is not a polished and finished product, but the result of an on-going effort of research and development.

Particularly, bugs in the software are still being discovered and corrected. Users of NS-2 are responsible for verifying for themselves that their simulations are not invalidated by bytes.

Similarly, users of NS-2 are entirely responsible for verifying for themselves that their simulations are not invalidated because the model implemented in the simulator is not the model that they were expecting a situation that does not exist with GNS-3.

OPNET – Optimized Network Engineering Tools

OPNET is a commercially available network engineering simulator marketed by OPNET technologies, INC it provides performance analysis for computer networks R and D (Research and Development), analysis and design of communication networks, devices, protocols and applications. OPNET alters users to analyze simulated networks to compare the impact of different technology designs on end-to-end behavior. OPNET incorporates a broad suite of protocols and technologies and includes a development environment to enable modeling of all network types and technologies including:

VOIP, TCP, OSPF V3, MPLS, IPV6

OPNET provides:

- One of the fastest event simulation engine among leading industry solutions
- Hundreds of protocol and render device models with source code
- Object-oriented modelling along with key features.

OPNET is not OPEN SOURCE software. However there is a free version named the IT GURU ACADEMIC VERSION which has very limited features including support for BGP-4 and Inter Domain Routing.

GNS-3 is preferred and actually used for the experiment/simulations because of the unreliability of NS-2 and the license requirements associated with OPNET. GNS-3 is also preferred because it is OPEN SOURCE and has excellent support and compatibility with both CISCO and JUNIPER equipment which most service provides employ to deliver end-to-end connectivity.

3.4.2 SCENERIOS

The purpose of this experiment is to demonstrate a Border Gateway Protocol (BGP) session hijack. To facilitate this demonstration, Graphical Network Simulator 3(GNS-3) running Cisco routers have been employed.

The experiment consists of three main scenarios;

i. Scenario one – this depicts a normal BGP operation.

ii. Scenario two – in this scenario the BGP session hijacking is demonstrated.

iii. Scenario three – the third and final scenario shows how this session hijacking can be prevented.

SCENARIO ONE - NORMAL OPERATION OF BGP

Routers R1 to 7 all represent routers in different Autonomous Systems (ASs), each having multiple point-to-point uplinks through other transit ASs to the global Internet. Each of the afore-mentioned routers is running external BGP (EBGP) and is peering with each directly connected router. The routers that are of particular interest are routers R6 and R7. For the purposes of this lab, R6 will serve as the AS originating a public Internet Protocol (IP) address, **20.20.0.0/19**, unto the global Internet. Inside AS 600, a server with a **20.20.20.20/24** address exists. All the routers in the various AS can access this server in AS 600. The BGP routes on each of the routers point to R6 to reach the 20.20.20.20/24 server. This is BGP in normal operation.

SCENARIO TWO - SESSION HIJACKING

The session hijacking comes into effect when router R7, begins to advertise a **20.20.20.0/21.** This is a subnet of the 20.20.0.0/19 being advertised by AS 600 (R6) albeit more specific prefix. The natural tendency of routing protocols to prefer a more specific route kicks in and the routers on the Internet (R1 through to R7) now use AS 700 (R7) as their path to reach the 20.20.20.20/24 IP.

This initially creates a 'black hole' on the Internet due to the fact that R7 does not actually have a node on its network with the 20.20.20.20/24 IP. To resolve this 'black hole' so R7 session hijacking can go unnoticed, R7 makes its advertisement/route for the 20.20.20.0/21 prefix undesirable to router, R4 and R6 by prepending their ASs (400 and 600) in its advertisement to the aforementioned routers. This ensures that routers R4 and R6 never use R7 as their primary path to the 20.20.0.0/19 network, but rather R6s path. R7 then uses a static route to point any traffic coming through it and destined for the 20.20.20.20/24server, to use R4 as its next hop router. This technique allows R7 to quietly receive all traffic meant for the 20.20.20.20/24 server, inspect and or alter it, and then forward it through R4 to its intended destination.

SCENARIO THREE - SOLUTION

To prevent this particular kind of BGP session hijacking, it is imperative that all the upstreams (typically ISPs) of the various ASes verify that their downlinks, i.e. the routers advertising routes through them, actually own the prefixes they are announcing. These uplinks must then setup filters to ensure that their downlinks are only allowed to advertise the routes that they own and nothing else.

In the case of this lab, R2, R8 and R4, the uplink providers of R7 check and implement filters to allow R7 to advertise only the routes that belong to that AS.

LAB TOPOLOGY

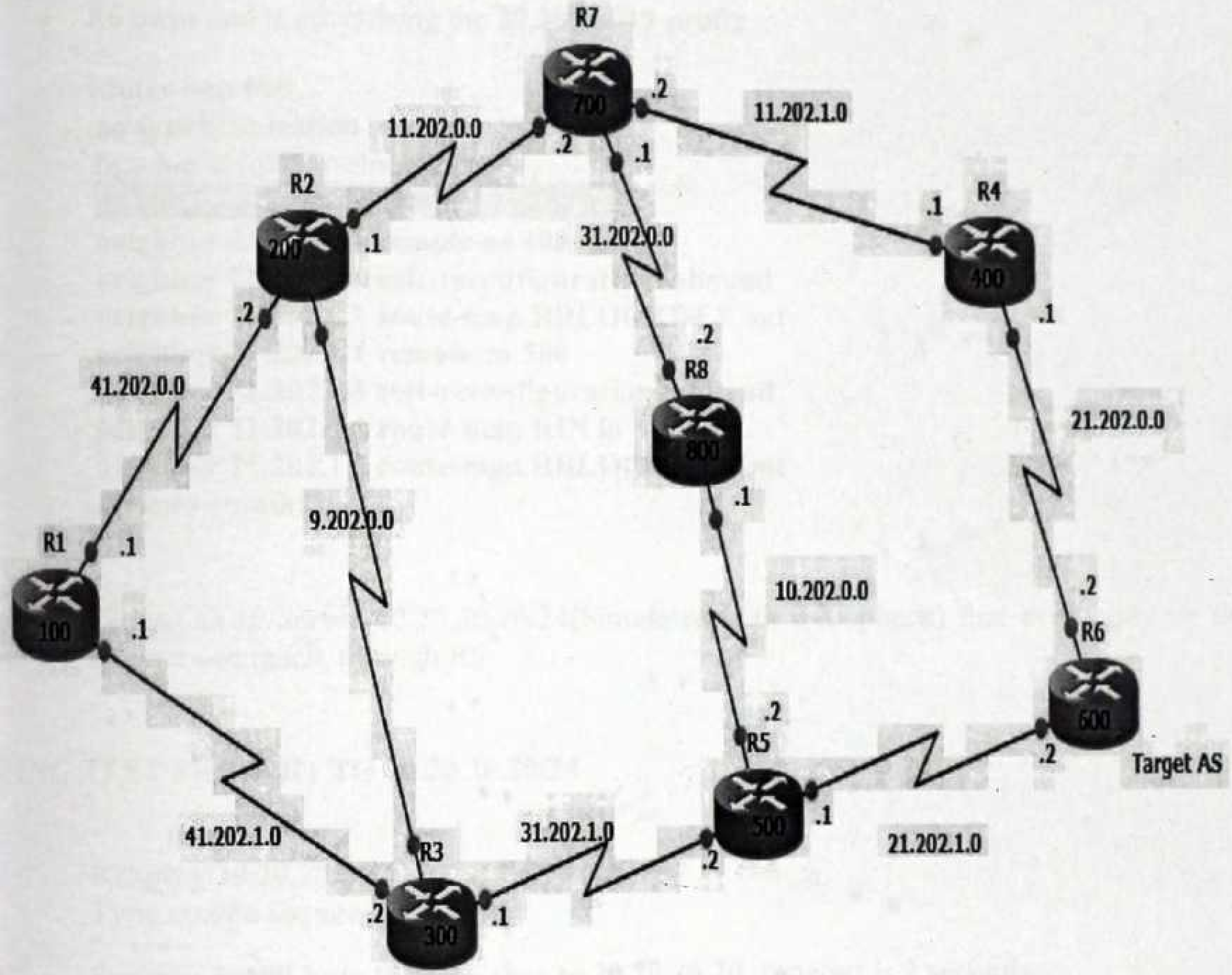

Figure 3.6 shows the topology for BGP experiment

3.4.3 SCENARIO ONE

- Normal BGP peering going on
- R6 in AS 600 has two uplinks to the global Internet, R4 and R5
- R6 owns and is advertising the **20.20.0.0/19 prefix**

```
router bgp 600
 no synchronization
 bgp log-neighbor-changes
 network 20.20.0.0 mask 255.255.224.0
 neighbor 21.202.0.1 remote-as 400
 neighbor 21.202.0.1 soft-reconfiguration inbound
 neighbor 21.202.0.1 route-map RBLOCKDEF out
 neighbor 21.202.1.1 remote-as 500
 neighbor 21.202.1.1 soft-reconfiguration inbound
 neighbor 21.202.1.1 route-map RIN in
 neighbor 21.202.1.1 route-map RBLOCKDEF out
 no auto-summary
```

- R6 has an IP/Server 20.20.20.20/24(Simulated with a loopback) that everybody on the Internet can reach, through R6

PING TEST FROM R1 TO 20.20.20.20/24

```
R1#ping 20.20.20.20
Type escape sequence to abort.

Sending 5, 100-byte ICMP Echos to 20.20.20.20, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 28/50/72 ms
```

TRACEROUTE TO 20.20.20.20/24 FROM R1

```
R1#traceroute 20.20.20.20

Type escape sequence to abort.

Tracing the route to 20.20.20.20
```

```
1 41.202.1.2 20 msec 4 msec 40 msec
2 31.202.1.2 [AS 300] 32 msec 40 msec 16 msec
21.202.1.2 [AS 500] 48 msec *  40 msec
```

PINGS TO 20.20.20.20 FROM R7

```
R7#ping 20.20.20.20

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 20.20.20.20, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 32/43/52 ms
```

TRACEROUTE FROM R7 TO 20.20.20.20

```
R7#traceroute 20.20.20.20

Type escape sequence to abort.
Tracing the route to 20.20.20.20

 1 11.202.1.1 4 msec 8 msec 28 msec
 2 21.202.0.2 [AS 400] 16 msec *  36 msec
```

- R1 and R7 are learning about the routes from BGP via both R6's uplinks thus R4 and R5 as shown in the traceroute. R1 uses R5 (AS 500) to get to the 20.20.20.20/24Server

TRACEROUTE TO 20.20.20.20/24 FROM R1 ENDS AT R6 21.202.1.2 IP

```
R1#traceroute 20.20.20.20

Type escape sequence to abort.
Tracing the route to 20.20.20.20

 1 41.202.1.2 20 msec 4 msec 40 msec
 2 31.202.1.2 [AS 300] 32 msec 40 msec 16 msec
 3 21.202.1.2 [AS 500] 48 msec *  40 msec
```

R7 uses R5 (AS 400) to get to the 20.20.20.20/24Server

TRACEROUTE FROM R7 TO 20.20.20.20 ENDS AT R6 21.202.0.2 IP

```
R7#traceroute 20.20.20.20

Type escape sequence to abort.
Tracing the route to 20.20.20.20

 1 11.202.1.1 4 msec 8 msec 28 msec
2 21.202.0.2 [AS 400] 16 msec *  36 msec
```

3.4.4 SCENARIO TWO

A Rogue Router, R7, in as 700 begins to advertise a more specific route (20.20.20.0/22) than the 20.20.0.0/19 being advertised by R6. Due to the nature of routing protocols preferring longer/more specific prefixes, all the BGP routers now point to AS 700 to reach the 20.20.20.20/24 server. R7 originally does not have a server with the specified IP so traffic meant for that server coming from the BGP routers all end at R7's AS 700 and get dropped, essentially creating a 'blackhole' for the 20.20.20.20/24 IP on the Internet. The 'blackhole' is rectified by using an AS prepend to make the R7s path to the 20.20.0.0/19 network and ultimately the 20.20.20.20/24 server undesirable.

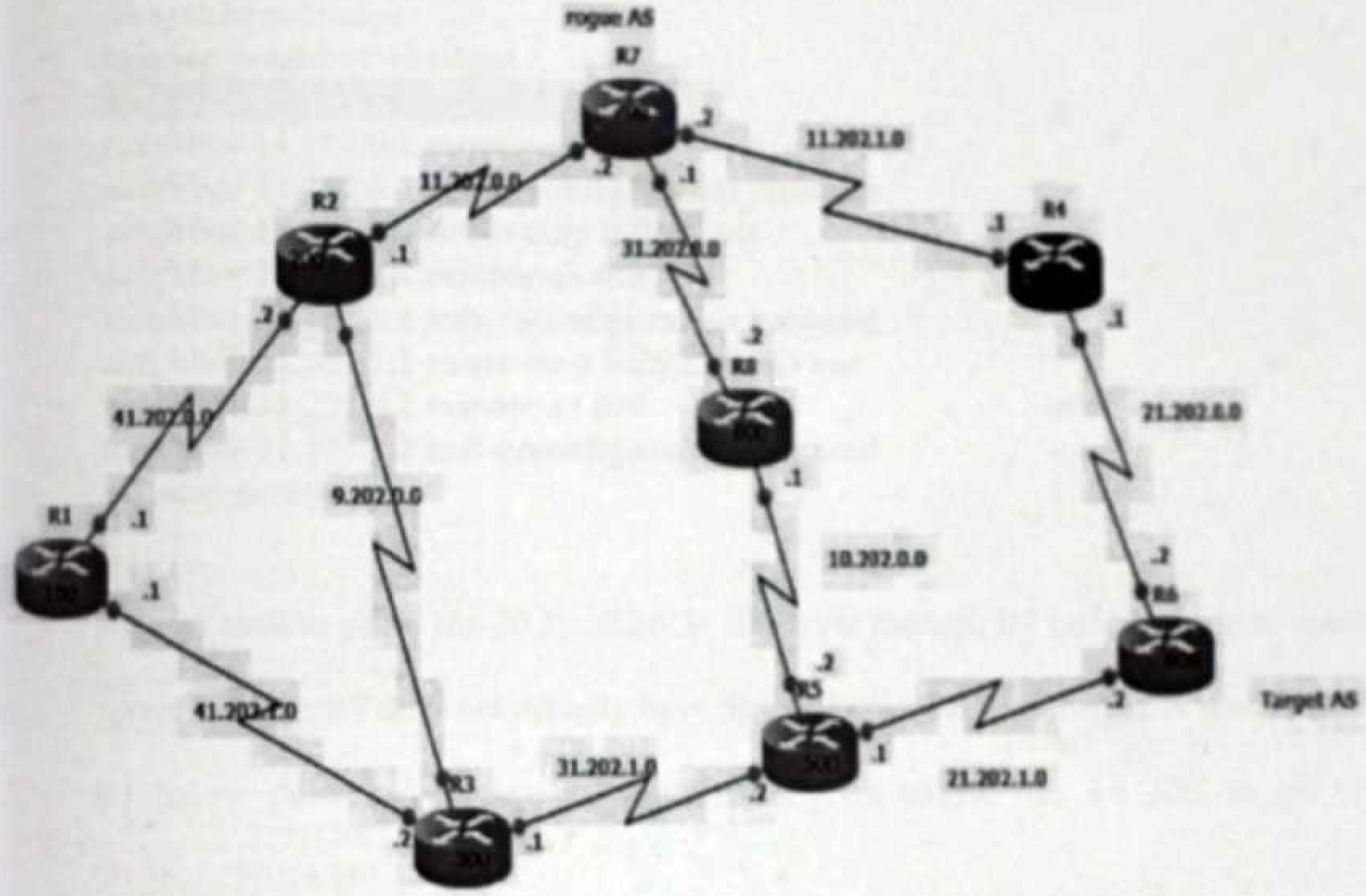

Figure 3.7 R7 begins to originate a more specific route of 20.20.20.0/22

```
R7#sh run | sec router bgp

 router bgp 700
 no synchronization
 bgp log-neighbor-changes
 network 20.20.20.0 mask 255.255.252.0
 neighbor 11.202.0.1 remote-as 200
 neighbor 11.202.0.1 soft-reconfiguration inbound
 neighbor 11.202.0.1 route-map RDEF out
 neighbor 11.202.1.1 remote-as 400
 neighbor 11.202.1.1 soft-reconfiguration inbound
 neighbor 11.202.1.1 route-map RPREPEND out
 neighbor 31.202.0.2 remote-as 800
 neighbor 31.202.0.2 soft-reconfiguration inbound
 no auto-summary
```

- R1 now tries to get to the 20.20.20.20/24 IP/Server through R7 but is unable to reach the server because R7 does not actually have that IP/Server on its network. A traceroute from R1 below shows R1 change its route to the R7's uplink R2, AS 200, to get to the 20.20.20.20/24 IP.

TRACEROUTE FROM R1 TO 20.20.20.20/24 ENDS AT R7 11.202.0.2 IP

```
R1#traceroute 20.20.20.20

Type escape sequence to abort.
Tracing the route to 20.20.20.20

  1 41.202.0.2 12 msec 8 msec 36 msec
  2 11.202.0.2 [AS 200] 92 msec 20 msec 24 msec
  3 11.202.0.2 [AS 200] !H  *  !H
```

- Ping test from R1 to the 20.20.20.20/24 ip reports the destination to be unreachable

```
R1#ping 20.20.20.20

Type escape sequence to abort.
```

```
Sending 5, 100-byte ICMP Echos to 20.20.20.20, timeout is 2 seconds:
U.U.U
Success rate is 0 percent (0/5)
```

- R7 resolves the blackhole by prepending the as 400 and 600 to its 20.20.20.0/22 advertisement to r4 and r6. This makes r7s path to the 20.20.0.0/19 network and ultimately the 20.20.20.20/24 server undesirable to r4 and r6.
- Using the **route-map** and **access control list prepend** and **accdef respectively**, r7 **prepends** the 400 and 600 ASes to its 20.20.20.0/22 advertisement to r4 and r6.

```
router bgp 700

 no synchronization
 bgp log-neighbor-changes
 network 20.20.20.0 mask 255.255.252.0
 neighbor 11.202.1.1 remote-as 400
 neighbor 11.202.1.1 soft-reconfiguration inbound
 neighbor 11.202.1.1 route-map RPREPEND out
 no auto-summary
!
ip access-list standard ACCDEF
 permit 20.20.20.0 0.0.3.255
!
!
route-map RPREPEND permit 10
 match ip address ACCDEF
 set as-path prepend 400 600
!
route-map RPREPEND permit 15
```

- R4 and R6s routing table now show R6s (21.202.0.2) as the best path to the 20.20.0.0/19 network

```
R4#sh ip bgp 20.20.0.0
BGP routing table entry for 20.20.0.0/19, version 3
Paths: (1 available, best #1, table Default-IP-Routing-Table)
Advertised to update-groups:
1
600, (received & used)
```

```
21.202.0.2 from 21.202.0.2 (20.20.20.20)
Origin IGP, metric 0, localpref 100, valid, external, best
```

```
R6#sh ip bgp 20.20.0.0

BGP routing table entry for 20.20.0.0/19, version 2
Paths: (1 available, best #1, table Default-IP-Routing-Table)
Advertised to update-groups:
1
Local
0.0.0.0 from 0.0.0.0 (20.20.20.20)
Origin IGP, metric 0, localpref 100, weight 32768, valid, sourced, local, best
```

- R7 now quietly redirects all traffic passing through it to the 20.20.20.20/24 server by employing a static route to point to r4 (11.202.0.2/30) as the next-hop to the 20.20.20.20/24 server.

STATIC ROUTE ON R7 POINTING TO R4

```
ip route 20.20.20.0 255.255.252.0 11.202.1.1
```

- R1 is now able to reach the 20.20.20.20/24 server through R7

Ping test from R1 to the 20.20.20.20/24 server

```
R1#ping 20.20.20.20

Type escape sequence to abort.
Sending 5, 100-byte ICMP Echos to 20.20.20.20, timeout is 2 seconds:
!!!!!
Success rate is 100 percent (5/5), round-trip min/avg/max = 40/62/108 ms
```

A traceroute from R1 to the 20.20.20.20/24 server

```
R1#trace 20.20.20.20
```

```
Type escape sequence to abort.
Tracing the route to 20.20.20.20

  1 41.202.0.2 24 msec 64 msec 8 msec
  2 11.202.0.2 [AS 200] 36 msec 28 msec 52 msec
  3 11.202.1.1 [AS 400] 48 msec 48 msec 40 msec
  4 21.202.0.2 [AS 400] 44 msec *  52 msec
```

3.4.5 SCENARIO THREE

- To prevent R7 from advertising a route/prefix that does not belong to that AS, the administrators of R7s uplink providers i.e. R4, R8 and R2 must verify from the internet registry which prefixes the AS700 can originate or announce. As a proactive measure, R4, R8 and R2 must also implement filtering mechanisms to ensure that R7 only advertises the routes that belong to it.

- In this lab, R7 owns and advertises the 70.70.70.0/24 prefix. After ensuring that R7 can and is allowed to advertise that prefix, R4, R8 and R2 implement the **route-map** below to filter R7s routes so it can only advertise the 70.70.70.0/24 route. However note that R7 can also choose to advertise a subset of the 70.70.70.0/24 prefix and its upstream providers can alter their filters to accommodate those prefixes.

A filter on R4, R8 and R2 to allow R7 to advertise only the 70.70.0.0/19 prefix

```
ip access-list standard PERMITVALIDIP

 permit 70.70.70.0 0.0.0.255
!
!
route-map PERMITVALIDROUTE permit 10
 match ip address PERMITVALIDIP
```

```
!
route-map PERMITVALIDROUTE deny 15
```

- R4, R8 and R2 now only receive the 70.70.70.0/24 advertisement from R7 preventing it from propagating the 'illegal' 20.20.20.0/22 route. This effectively stops the BGP session highjack.

The only valid route R4 receives from R7 is the 70.70.0.0/19 prefix

```
R4(config)#do sh ip bgp nei 11.202.1.2 received-route
BGP table version is 52, local router ID is 31.202.0.2
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal,
          r RIB-failure, S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network          Next Hop            Metric LocPrf Weight Path
*> 70.70.70.0/24    11.202.1.2               0             0 700 i
```

- R1 goes back to its original path to reach the 20.20.20.20/24 ip.

```
R1#trace 20.20.20.20

Type escape sequence to abort.
Tracing the route to 20.20.20.20

  1 41.202.1.2 28 msec 96 msec 28 msec
  2 31.202.1.2 [AS 300] 8 msec 56 msec 28 msec
  3 21.202.1.2 [AS 500] 84 msec *  96 msec
```

*compare with traceroute of scenario one.

3.4.6 CONCLUSION

The objective of the above BGP session hijacking simulation is to bring to light the inherent vulnerability of External BGP (EBGP) sessions and the measures taken to mitigate this vulnerability. A typical EBGP peering between different ASes on the Internet is described, a setting in which a number of ASes serve as uplinks/upstreams, passing on routes from the ASes originating these routes from their local systems, unto the global Internet.

In each of the scenarios, the flexibility of BGP is demonstrated. Naturally, BGPs comprehensive stock of attributes allows for routes to be altered, forwarded and dropped with relatively minimal fuss. The widely accepted notion that BGP is and can be tweaked to behave in many ways and choose any of several paths as it traverses the Internet makes it difficult to pinpoint a particular behavior as malicious.

For example, in the above scenarios, R2 could modify the routes it receives from R1, its downstream peer to make them seems like R2 was originating those routes. This might not necessarily constitute malicious behavior since R2 could do this to make some routes seem preferable to other routes it receives from other peers.

Again Router, R3 could for instance drop routes originating from R1's AS, but forward every other route. BGPs nature therefore makes policing the protocol to identify inconsistent behavior not entirely straightforward.

Concerns have been raised that perhaps ~~it is~~ about time a different Inter-domain routing protocol with fewer 'quirks' and a more well defined stable set of rules (in a security sense) be developed to replace the already ubiquitous BGP. This argument might hold some merit considering the fact that BGPs flexibility is what allows for our attacker (R7) to surreptitiously manipulate routes in such a way as to suite its malicious intentions.

Verifying and filtering the routes advertised by immediate peers can go some way to make BGP a more secure Inter-domain routing protocol. However, this might be an exercise in futility unless all peers thereof participate in this process of verification and filtering of routes. The task of getting every provider to filter routes is a daunting if not seemingly improbable quest.

In conclusion, it is imperative to point out that although BGP might be inherently flawed mainly because the propagation of routes by peering ASes is based fundamentally on trust, and this undoubtedly raises several levels of security concerns, however, the tradeoff between security and BGPs presently unmatched ability to adapt and traverse myriad paths to re-establish lost connectivity cannot be ignored.

CHAPTER 4 BGP MISCONFIGURATION

4.0 INTRODUCTION

BGP is important for the general reliability of the internet as it is the only inter-domain routing protocol in use as of now. There have been disruptions in some regions of the internet arising out of configuration errors (misconfiguration) in the BGP implementation of ISPs (Deshpande et al, 2008).

Although there are a number of known sources of misconfiguration, a definite mechanism or tools is not entirely available for the identification and quantification of BGP misconfiguration. This section presents perspectives on misconfiguration formed on the basis of secondary data made available online at Route-Views. The views formed and expressed here in this chapter are largely reliant on data and the methods for gathering the hence any assumptions made by the data source influence this analysis.

The objective as must be stated is not necessarily to generate data as to the frequency of BGP misconfiguration but rather to shed more light on the causes, the impacts on the global routing infrastructure as well as the policy implementations or strategies service providers and ASes in general can adopt to mitigate against misconfiguration with little or no changes to the existing routing equipment.

4.1 MISCONFIGURATION

BGP Misconfiguration can be defined as configuration errors that result in the unintended production or suppression of BGP routing announcements (Deshpande et al, 2008)

A misconfiguration encompasses both unavoidable errors and human errors arising out of human factors (Reason, 1990). The analysis of BGP misconfiguration requires an individual verification of every suspected anomaly in order to differentiate misconfiguration in the global routing tables from other forms of problems (part of the methodology used for data collection).

Although there could be several other flavors of misconfiguration, the focus remains on export and origin misconfiguration as stated in earlier chapters.

4.1.1 ORIGIN MISCONFIGURATION

An origin misconfiguration is said to have occurred when a service provider or an autonomous system (AS) for that matter is seen to have accidentally injected a prefix into the global BGP routing tables (McDaniel, 2005).

This could be the result of failure in summarizing an address space which leads to the insertion or injection of a number of specific routes which ideally should not have been advertised. This could also occur as a result of session hijacking which results in the announcement of part of other ISPs or AS's address space (Reason, 1990).

Origin misconfiguration could also occur when routes meant to be private thus only advertised within a private network or IGP, are propagated beyond the border routers of the private network.

4.1.2 EXPORT MISCONFIGURATION

Export misconfiguration could be better explained with a scenario as below:

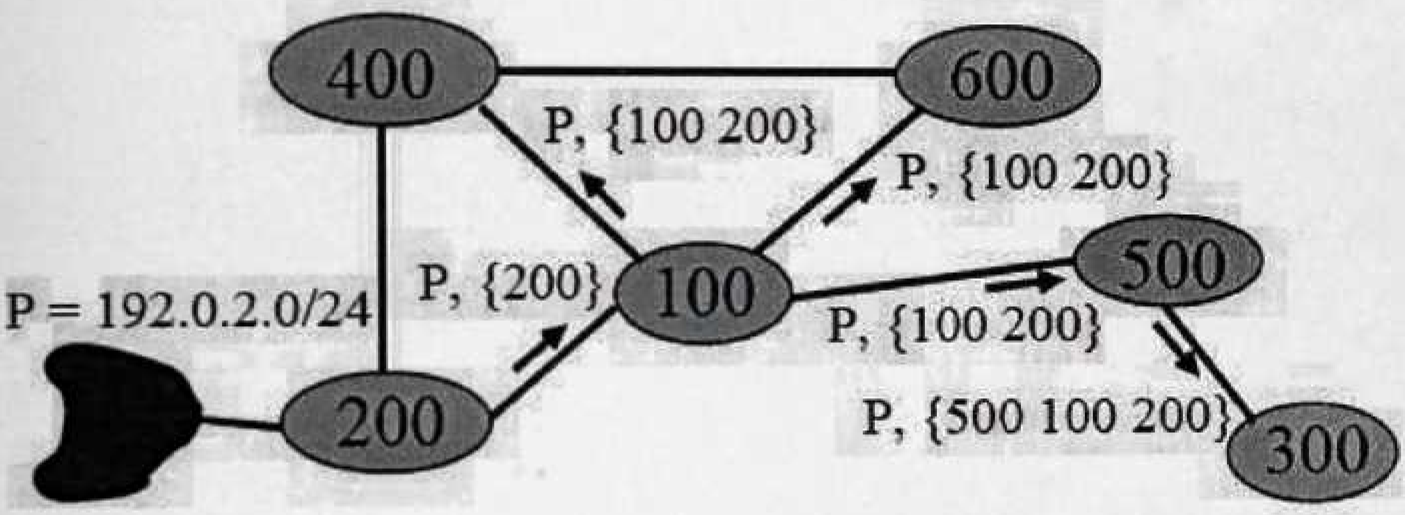

Figure 4.1 export misconfiguration occurs which AS400 exports route P to AS600
Source: Mahajan et al, 2001

An export misconfiguration could be said to have occurred if the AS-path happens to be in violation of the policies of one of the ASes in its path simply because the router exported a route if should have filtered.

Using the diagram above, an export misconfiguration could be said to have occurred if for example AS400 exports the route P to AS600 against its policy. Of the numerous other problems identified with BGP, misconfigurations have been chosen for study largely due to its potential to widely disrupt internet connectivity.

As part of misconfiguration in general, an AS for instance may accidentally filter out routes it otherwise meant to announce to a remote observer.

4.2 EFFECTS OF MISCONFIGURATIONS

BGP misconfigurations in certain cases are unintentional however their impact can be felt negatively regardless the kind of error or misconfiguration that occurs. Some of the negative effects on ASes arising out of misconfiguration include:

i. EFFECT ON ROUTING LOAD

BGP misconfiguration may increase the routing load by generating unnecessary BGP updates. A lot more BGP routers are already heavily loaded due to the rapid growth of the internet and its routing registries (Masan, 2001).

Any unnecessary overhead therefore could result in degraded services and poor response time which is a source of concern in the BGP operations arena.

ii. INTERRUPTION OF CONNECTION.

Connectivity could be affected as a result of misconfiguration; either partially thus only affecting some parts of the internet or globally thus from everywhere on the entire internet.

iii. VIOLATION OF ROUTING POLICIES

As per the definition, BGP misconfiguration could potentially violate the policies intended for a particular Autonomous System. Prefixes could be leaked incorrectly to the entire internet and as a result, the erroneous routes announced may be preferred by ASes over more legitimate ones, causing an AS to be used as a transit AS by other ASes for which it had no previous transit arrangements with.

4.3 ORIGIN MISCONFIGURATION ANALYSIS

When a route is inserted accidentally or unintentionally into the global BGP tables, an origin misconfiguration could be said to have occurred. This kind of error could show up as a short-lived new route assuming the ISP or AS realized the mistake and quickly rectified it.

To ascertain the validity of this assertion as to whether the event is actually a misconfiguration, historical BGP data from the previous day is used to differentiate between new routes and old routes that reappear due to the ending of a failure period. Subsequently BGP updates are then used to determine how long a route lasted. New routes that appear because of policy changes such as multi homing, traffic engineering and provider switch expected to last longer than those due to misconfiguration.

For a better understanding on the causes of misconfiguration, the new routes are classified based on their relationship with the existing routes as shown in the Table 4.1 below.

Table 4.1 shows Export policies for common commercial relationships.

Route export	Export Policy
Customer »»»Provider	Only routes received from customers and siblings.
Peer »»» Peer	Only routes received from customers and siblings
Provider »»» Customer	All routes
Sibling »»» sibling	All routes

As shown in table 4.1;

- Providers provide transit to customers
- Peers exchange only traffic that is sourced by them, their customers or their siblings, and siblings provide mutual transit.

For each category, one of the possible new routes is listed along with the old route. In self deaggregation which is the de-summarization of a received prefixes received by an AS or origin, the origin deaggregates its own prefix. (Thus expands a prefix into a more specific route).

In related origin, an existing prefix or its subset is announced by a new origin that appears related to the old origin in that one of the origins appears in the AS-path of the other.

A prefix or its subset is advertised by a different origin in foreign and the two origins apparently have no relationship to one another. New routes for prefixes that are neither present in the table nor have a less specific prefix in the table are also classified as a foreign origin incident.

The classification is done along the lines of the most likely underlying causes. Self deaggregation incidents for instance could be the result of forgetting to aggregate (summarize) a route at a router, while foreign origin incidents could be the result of an address space hijacking.

An origin that is related could be actually connected to the network than a foreign origin is, though the latter can also be caused by a backup origin that only appears during failures.

4.4 EXPORT MISCONFIGURATION ANALYSIS

An export misconfiguration could be said to have occurred as a result of an inadvertent export of a route to a BGP peer which violates the exporting AS's routing policies (Mahajan et al, 2001)

For commercial reasons, AS relationships are closely guarded secrets which make the process of identifying such export misconfigurations a more cumbersome one.

(GAO, 2000) infers these relationships from the BGP tables based on a number of observations.

Assumption 1: All valid AS-path are valley free treating the provider to customer direction as downward and siblings and peers at the same level, the valley free property means that a route that starts going downwards never goes up again.

Assumtion2: AS-path can have at most one peer-to-peer edge which occurs at the highest point in the path.

Assumption 3: ASes with more neighbors are more likely to be providers.

Based on the relationships inferred by Gao's algorithm (Gao, 2000), and using historical BGP data as input, AS-paths with short-lived sub paths that violate the valley free condition or contain multiple peering edges identified as possible export misconfigurations.

The inference of AS relationships is not an all round perfect approach and as a result real misconfigurations could be omitted and vice-versa.

Export misconfigurations could be categorized as shown in table 4.2

Table 4.2 Classification of Route Export misconfigurations

Export	Policy Violation
Provider-AS-Provider	Route exported to provider was imported from a provider.
Provider-AS-Peer	Route exported to peer was imported from a provider
Peer-AS-Provider	Route exported to provider was imported from a peer.
Peer-AS-Peer	Route exported to peer was imported from a peer.

As shown in table 4.2 above, the AS provides transit to traffic from its provider or peer to its provider or peer in each instance. The classification is based on the policy being violated. A route export misconfiguration could also contain siblings, for simplicity, a chain of siblings is considered to be one AS.

4.5 CAUSES OF MISCONFIGURATIONS

To prevent misconfigurations, a deeper understanding as to why they occur is necessary. This sub section identifies and discusses the causes of the misconfigurations. A standard classification of human errors helps to categorize the causes of BGP misconfiguration (Mahajan et al, 2001).

The broad categories of human errors include:

I. **SLIPS:** slips refer to errors in the execution of an otherwise correct plan. (Deshpande et al, 2008). This category may include typographical errors and the accidental omission of a configuration command.

ii. **MISTAKES:** mistakes refer to errors where the execution was carried out as planned but the plan itself was incorrect form the very onset. Logical bugs and improper operational practices could be cited as examples.

It must be stated here that there is not much clarity in the distinction between a slip and an error hence in instances of doubt; the cause is classified as a slip with the assumption that the operator might have constructed an accurate mental plan for the configuration change.

4.5.1 CAUSES OF ORIGIN MISCONFIGURATION

The causes of BGP origin misconfiguration are broadly attributed to two main factors namely

I. Software initialization bugs

ii. CLI (command line interfaces)

Table 4.3 BGP misconfiguration data
Source: Route-Views Server

Misconfigurations	*Prefixes*		*Incidents*		*Type*
	Total	Connectivity	Total	Connectivity	
Initialization bug	1580 (22%)	0 (00%)	43 (05%)	0 (00%)	mistake
Reliance on upstream filtering	977 (14%)	0 (00%)	431 (46%)	0 (00%)	mistake
Old configuration	72 (01%)	28 (39%)	36 (04%)	20 (56%)	mistake
Redistribution	2294 (32%)	1 (00%)	43 (05%)	1 (02%)	slip
Community	99 (01%)	2 (02%)	28 (03%)	2 (07%)	slip
Hijack	101 (01%)	101 (100%)	54 (06%)	54 (100%)	slip
Forgotten filter	53 (01%)	1 (02%)	13 (01%)	1 (08%)	slip
Incorrect summary	26 (00%)	0 (00%)	17 (02%)	0 (00%)	slip
Unknown configuration error	1053 (15%)	39 (04%)	90 (10%)	12 (13%)	slip
Miscellaneous	88 (01%)	16 (18%)	38 (04%)	10 (26%)	
Unclassified	779 (11%)	82 (11%)	152 (16%)	23 (15%)	

Non-misconfigurations	*Prefixes*	*Incidents*	
Failure	91 (34%)	50 (32%)	
Testing	66 (24%)	44 (28%)	
Migration	51 (19%)	26 (17%)	
Load balancing	22 (08%)	20 (13%)	
Miscellaneous	11 (04%)	7 (04%)	
Unclassified	30 (11%)	10 (06%)	

INITIALIZATION BUGS

The data made available for this analysis suggests a certain frequency of occurrence in which ASes would announce a lot more specific prefixes which get withdrawn within a few minutes. Some of the reasons given include unintentional reboot of their edge routers; a flapping interface or the running of maintenance scripts.

While a router is rebooting or the filters being updated the more-specific prefixes present in the router's table can be leaked only to be withdrawn when the reboot or update is complete.

Upon further researching, these instances were identified to be possible error in the operating software of some routers (IOS in the case of CISCO) or the interaction of configuration processing semantics with the way operators write their maintenance scripts or both.

RELIANCE ON UPSTREAM FILTERING

Some autonomous systems announce routes based on the assumption that they would be filtered by their upstream provider, and hence might not become globally visible. In certain instances, prefixes were injected and withdrawn within a short time period.

The operators involved in these situations later realized that their announcements were seen beyond their immediate upstream neighbors and had to take steps to have the problem rectified by their providers.

OLD CONFIGURATIONS

Another cause of BGP UPDATE misconfiguration that led to connectivity problems is the problem of old configuration. In certain instances, some operators had the configuration on their routers changed correctly but did not commit the changes they made to a stable state or storage.

A separate command is required on most routers however there is no warning about unsaved configurations hence when the router reboots the next time; the old configuration comes back into effect. The router then begins to announce old routers.

REDISTRIBUTION

Redistribution allows an operator to specify which routes learned from other routing protocols (e.g. IGP's) such as OSPF, should be advertised to BGP peers. There are a number of ways by which redistribution could generate errors however the two subsequent instances would be used to analyze the problem.

REDISTRIBUTE

Redistribute igrp 100 route-map igrp2bgp.

This line of command, tells the router not to advertise everything in IGP (Interior Gateway Protocol) tables that matches the route map is unable to specify the route-map part or gets the route-map itself wrong, all the prefixes in the IGP tables would be announced via BGP.

```
*aggregate-address                    summary-only

Aggregate-address                     192.168.0.01255.255.0.0
```

This line of command, instructs the router not to advertise any subsets of the prefix. However if summary-only is forgotten, all the more-specific prefixes in the routing table would be advertised.

If redistribution is not done appropriately, it could lead to a large number of faulty prefix advertisements such as the AS7007 incident (Misel, 1997) which would expose BGP to variations in the IGP protocol.

COMMUNITIES

Another major cause of origin misconfiguration is the misapplication of the community attribute of BGP. Attaching the wrong community attribute to a prefix was found to be a major contributor to the problem.

The community attribute is utilized by Autonomous systems to mark their routers and also to relay policies such as “don’t propagate further” or “export only to your immediate peers”.

Prefixes potentially could be propagated beyond where they were originally intended or in some instances the prefixes might not got propagated at all.

HIJACKS

When an unrelated Autonomous System (AS) announces a prefix or an address space (e.g. K-NET advertising the 41.202.0.0/19 blocks belonging to BBH), belonging to another AS either as a result of a security flaw or as a result of a typographical error, origin misconfiguration could occur.

INCORRECT SUMMARY

The application of an incorrect summary mask can cause an AS to announce an ADDRESS block that might be larger or seemingly smaller than the actual block. For instance 255.255.0.0 is the right prefix mask for 192.168.0.40/16 but using 255.255.255.0 yields 192.168.0.0/24 and using ~~255.0.0~~.0 yields 192.0.0.0/8 resulting in incidents where smaller ASes have announced/8s that have not been allocated to any organizations as a single block according to the internet routing registries.

4.5.2 CAUSES OF EXPORT MISCONFIGURATIONS

This subsection (examines) attempts to elicit the causes of export misconfigurations based the data provided in table 4.4:

Table 4.4 snapshot BGP export misconfiguration data
Source: Route-Views Server

Misconfigurations	*Paths*	*Incidents*	*Type*
Prefix based config	98 (08%)	46 (22%)	mistake
Old configuration	20 (02%)	9 (04%)	mistake
Initialization bug	18 (01%)	9 (04%)	mistake
Bad ACL or route map	445 (34%)	8 (04%)	slip
Typo	153 (12%)	13 (06%)	slip
Forgotten filter	109 (08%)	15 (07%)	slip
Community	69 (05%)	37 (18%)	slip
Unknown config error	193 (15%)	14 (07%)	slip
Miscellaneous	22 (02%)	5 (02%)	
Unclassified	162 (13%)	54 (26%)	
Non-misconfigurations	*Paths*	*Incidents*	
Backup Arrangement	22 (47%)	13 (38%)	
Special Arrangement	13 (28%)	11 (32%)	
Failure	9 (19%)	7 (21%)	
Unclassified	3 (06%)	3 (09%)	

PREFIX BASED CONFIGURATION

Prefix based configuration is one of the major contributors to the export misconfiguration problem. This particular factor can be better explained with the diagram and scenario shown in figure 4.2

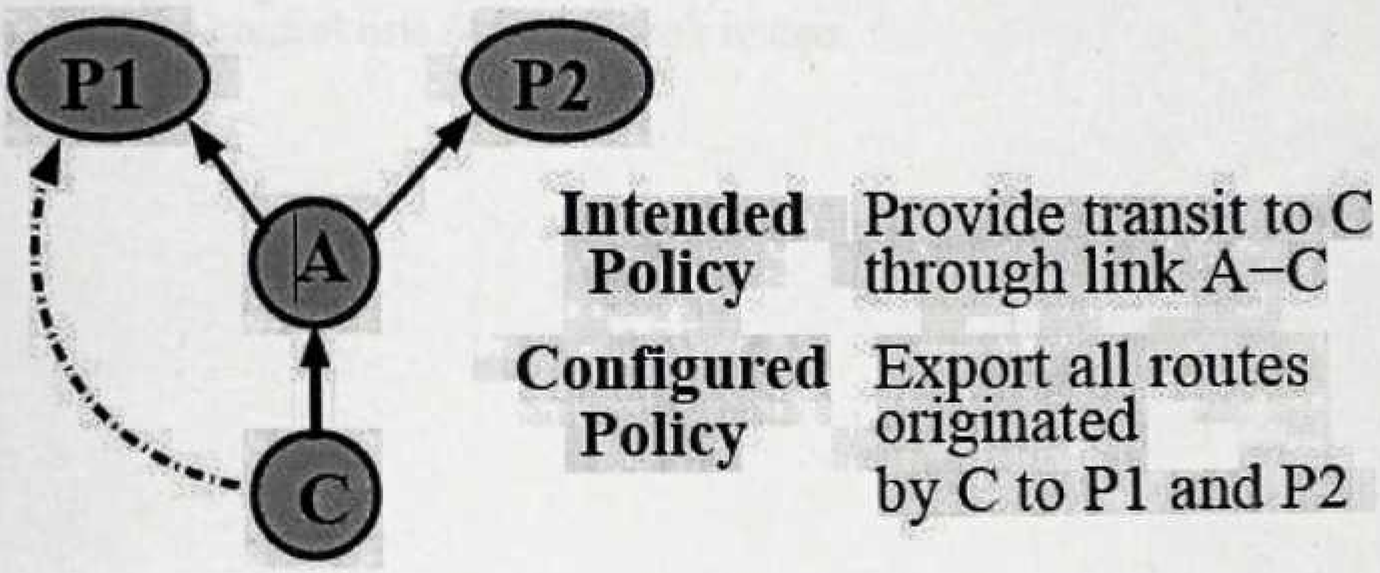
P1
P2
A
C
Intended Policy
Provide transit to C through link A–C
Configured Policy
Export all routes originated by C to P1 and P2

Figure 4.2 export misconfiguration occurs which occurs when link A-C fails
Source: Rexford et al, 2005

P1 and P2 are providers of A, and C is a customer of A. C also has a path to P1 that does not go through A. since A is the provider of C, it provides transit to C by exploiting routes to both P1 and P2. This policy is expressed as "it is ok to export C's prefixes to P1 and P2 which can be implemented by listing the prefixes explicitly or specifying that the origin AS should be C.

The configuration above works well when there are no failures. A always chooses a direct path to C hence all the traffic that comes to A meant for C is seat via link A-C. In the event of a failure ~~of the link~~ A-C, A would hear announcements for C's prefixes through P1. Based on the configuration, it would announce this route to P2, thus becoming a transit between P1 and P2 for all traffic for C. This misconfiguration can be prevented by configuring a different policy: export routes for C only when the AS-path is C.

BAD ACCESS CONTROL LIST OR ROUTE MAP

This situation applies to instances where operators make mistakes during the configuration of route maps or access control lists (ACL) on their routers.

4.6 DISCUSSION

BGP misconfiguration potentially could increase the internet's vulnerability to accidental errors if equipment manufacturers, vendors and service providers fail to advance technology and policies that attempt to address the situation.

This section examines counter measures to problem of misconfigurations. The suggestions are made based on the probable causes of misconfigurations and the secondary data made available in the previous subsections. A number of suggestions that could help forestall misconfiguration events include but is not restricted to the under listed.

4.6.1 IMPROVED USER INTERFACE DESIGN

Graphical user interfaces could be introduced on most industrial routers to further reduce the possibility of BGP misconfigurations. Graphical user interfaces would provide safe defaults and some consistency across the various router versions as well as the minimization of the dependence between multiple lines of configuration. CISCO systems have started in this direction with the introduction of the CISCO CONFIGURATION PROFESSIONAL GUI based router configuration and management platform. Other vendors such as JUNIPER, HUAWEI, HP and should be encouraged to introduced GUI's on their high end industrial routers as well.

It is clear from the causes of misconfigurations that of the benefits of GUI are not identifiable in the current command line interfaces (CLI's) currently provided on most industrial routers provided by the major vendors. However the introduction of GUI's could reduce the risk associated with the configuration of sensitive features such as redistribution by explicitly listing the prefixes to be announced through BGP rather than having to manually include the prefix as is the case with CLI (command line interface).

4.6.2. HIGH-LEVEL LANGUAGES AND CONFIGURATION CHECKING/AUDIT.

Network engineers build router configurations by entering the commands one line at a time. This is typically a low level task which is error-prone.

Configuration tools that could allow operators and network engineers express routing policies in a high-level form can reduce configuration errors that translate into misconfigurations.

Configuration checkers are also another consideration with respect to BGP misconfigurations. Configuration checker looks for consistency within and between configurations and warns the operators when a safety property has been violated e.g. when a required route map remains undefined.

It may also prove effective to annotate configurations with optional, high-level, declarative expressions of the intended policy as this would provide greater scope for consistency checking. CISCO systems as part of its CISCO CONFIGURATION PROFESSIONAL PLATFORM (Cisco Press, 2012) introduces an auditing and security management facility which should be adopted and encouraged to become industry standard across all other platforms regardless of the vendor.

4.6.3 DATABASE CONSISTENCY AND REGISTRIES

Misconfigurations could also occur as a result of incorrect or inconsistent data used for active configuration. There are a number of databases that service providers may have to refer to at a point when configuring their border routers to do BGP.

First, each router contains its own version of configuration information. Secondly, the service provider also maintains configuration related data in its network management system. Thirdly, allocation (thus IP address allocations) and routing policies may be maintained in industry-wide IRRs (Internet Routing Registries).

These databases necessarily have to be derived from each other to avoid inconsistency and errors. A possible resolution could be to build consistency mechanisms into routers e.g. allowing the router to update routing registry directly or at least checking the registry for consistency. There are no such checks and as a result, the information in the registries is largely believed to be inaccurate.

CHAPTER 5 CONCLUSION

The border gateway protocol currently used to exchange reachability information between Autonomous Systems over the internet has certain vulnerabilities which have been exploited to cause a session hijacking attack as demonstrated through the simulations in chapter three.

This possibility (thus the ability to launch a session hijacking/black hole) is occasioned by the absence of a formidable overall security implementation or framework despite the countless number of approaches, tools and protocol extensions put forward for the purpose of securing BGP. Simple misconfiguration have played a role in the security breached associated with BGP over the years notable among them being the AS7007 (Butler et al, 2006) incident of Pakistani Telecom.

BGP has also been susceptible to attacks as demonstrated in previous literature as a result of its own protocol design. BGP is based on trust and has no built in peer authentication mechanisms and no verification of update message; a situation that the protocol extension SBGP (secure BGP) attempts to address.

BGP uses TCP as a transport layer protocol to establish connections and exchange reach ability information with other border routers in other ASes. BGP becomes liable to all the vulnerabilities of TCP as a result of this dependence. This paper examined a number of techniques adopted for TCP security as well.

TCP MD5 encryption (Kent et al, 2005) and peer authentication is one measure adopted for protecting BGP from the weakness of its transport partner i.e. TCP. MD5 only provides peer authentication leaving the question on update message authentication still unanswered. Many other protocol extensions such as pretty secure BGP (psBGP), secure origin BGP (soBGP) and a host of others have been examined in this work with emphasis on the reasons why they are still not widely used by commercial internet service providers. The findings in this regard revealed that many of these relatively newer secure forms of BGP are cost ineffective as a result of their huge overhead and enormous demand on processing power making them commercially unattractive.

Another reason identified in this thesis is the negative effect of the overhead on the performance of BGP routers themselves that leads to degraded performance and slower BGP convergence times.

Routers running BGP were found to be already heavily loaded and could not bear the extra load of these heavy new protocols hence the reluctance of industry players in adopting and deploying these protocols.

This works takes a preventive approach in addressing BGP security vulnerabilities by proposing a security policy aimed at providing service providers with alternative mechanisms for securing BGP sessions and routers. This solution thrives on the implementation of already well known best practices and security measures available for BGP security. Specifically, this approach aims at making BGP routers more secure by allowing only restricted or controlled access as a first step towards security. Other mechanisms such as route filtering, route maps, policy maps are also to be configured and deployed for the same purpose. BGP routers must essentially be properly configured to prevent misconfiguration that could be exploited for session hijacking and other forms of attacks.

To ensure this (thus prevent session hijacking), all upstream neighbors in any BGP peering session which (might typically be ISPs); should have their edge routers configured to verify all routes advertised to them by their downstream neighbors. By verification; upstream neighbors should check that their downlinks actually own the prefixes they are announcing. The uplinks must then set up filters to ensure that their downlinks are only allowed to advertise routes that they own and nothing else.

This approach is clearly demonstrated with a three phase simulation based scenario in chapter 3. Scenario 1 begins by simulating the normal operation of BGP with a number of ASes connected together to mimic as sub section of the internet.

The results of the simulation confirmed the operational viability of the Security Policy proposed in chapter 3 and asserts its potential for increasing an ISPs chances of preventing session hijacking attacks despite the cumbersome configuration task involved.

It must be said that the Security Policy is not an ultimate solution to the problem as new forms of attacks and threats emerge over the internet consistently.

5.1 RECOMMENDATIONS

The internet is an indispensible tool or platform which needs to be constantly improved for the benefit of its numerous categories of users. In this regard, a lot more research is required to unearth the cutting edge innovations of the next generation internet.

Networks need to expand in terms of capacity and bandwidth. Reliability and security should also be guaranteed with the latter resting mostly on the implementation of more secure approaches to BGP security.

By way of recommendations:

- ISPs should ensure restricted and controlled physical access to all border routers
- Service providers should also engage in training and retraining for their engineers and technicians to bring them up to speed on the state of the art with respect to BGP configuration and internet routing in general.
- Only the very experienced engineers should be given access to manage and configure provider edge routers for ISPs.
- As a means of arresting the problem of misconfiguration, network equipment vendors and engineers should collectively drive the agenda for improvement in the CLI(command line interface) used for router configuration
- Where available, network engineers should utilize the GUI based router configuration and management platform such as the CISCO CCP (Cisco configuration professional) to prevent accidental misconfiguration.

5.2 FUTURE WORK

BGP has proved resilient in terms of its longevity of useful life despite early predictions and warning signals.

Future research could be targeted at the development and rigorous testing of GUI based configuration tools to mitigate the impact and occurrence of misconfiguration

Towards the eminent necessity for network convergence, research in the area of internet research should also pay some attention to high speed WAN transmission such as MPLS (Multi Protocol Label Switching) with particular emphasis on QoS (Quality of Service).

REFERENCES

1. I. Peter, "Origins of The Internet," [Online]. Accessed 2012, January 8. http://www.nethistory.info/History%20of%20the%20Internet/origins.html

2. M. Sachdeva, G. Singh, and K. Kumar, "Deployment of Distributed Defence against DDoS Attacks in ISP Domain," *International Journal of Computer Applications (0975 – 8887)*, Volume 15, Issue 2, pp. 29, February, 2011

3. Y. Rekhter, T. Li, "A Border Gateway Protocol 4 (BGP-4)" , RFC 1771, the Internet Engineering Task Force (IETF), 1995 [Online] Available at http://www.ietf.org/rfc/rfc1771.txt

4. Y. Rekhter, T. Li, "A Border Gateway Protocol 4 (BGP-4)" , RFC 4271, the Internet Engineering Task Force (IETF), 2006 [Online] Available at http://www6.ietf.org/rfc/rfc4271

5. M. Caesar and J. Rexford, "BGP Routing Policies in ISP Networks," IEEE Network, November/December 2005

6. M. Zhao, S. Smith, and D. Nicol, "The performance impact of BGP security," Network, IEEE, vol. 19, no. 6, pp. 42–48, Nov.-Dec. 2005.

7. R. Mahajan, D. Wetherall, and T. Anderson, "Understanding BGP misconfiguration," in SIGCOMM '02: Proceedings of the 2002 conference on Applications, technologies, architectures, and protocols for computer communications. New York, NY, USA: ACM, 2002, pp 2.

8. Office of the President of the United States, "Priority II: A National Cyberspace Security Threat and Vulnerability Reduction Program," 2004. [Online]. Available: http://www.us-cert.gov/reading room/cyberspace strategy.pdf

9. S. Murphy, "BGP Security Vulnerabilities Analysis," RFC 4272 (Informational), Internet Engineering Task Force, Jan. 2006. [Online] Available: http://www.ietf.org/rfc/rfc4272.txt

10. A. Barbir, S. Murphy, and Y. Yang, "Generic Threats to Routing Protocols," RFC 4593 (Informational), Internet Engineering Task Force, Oct. 2006. [Online] Available: http://www.ietf.org/rfc/rfc4593.txt

11. H. Ballani, P. Francis, and X. Zhang, "A study of prefix hijacking and interception in the Internet," SIGCOMM Comput. Commun. Rev., vol. 37, no. 4, pp. 265–276, 2007.

12. K. Zetter, "Revealed: The Internet's Biggest Security Hole," Wired Magazine - ThreadLevel, Aug 2008. [Online]. Available: http://www.wired.com/threatlevel/2008/08/revealed-the-in/

13. T. Underwood, "Con-Ed Steals the 'Net'," Renesys Blog, Jan 2006. [Online]. Available: http://www.renesys.com/blog/2006/01/coned-steals-the-net.shtml

14. M. A. Brown, "Pakistan hijacks YouTube," Renesys Blog, Feb 2008. [Online]. Available: http://www.renesys.com/blog/2008/02/pakistan-hijacks-youtube-1.shtml

15. V. Gill, J. Heasley, and D. Meyer, "The Generalized TTL Security Mechanism (GTSM)," RFC 3682 (Experimental), Internet Engineering Task Force, Feb. 2004, obsoleted by RFC 5082. [Online]. Available: http://www.ietf.org/rfc/rfc3682.txt

16. V. Gill, J. Heasley, D. Meyer, P. Savola, and C. Pignataro, "The Generalized TTL Security Mechanism (GTSM)," RFC 5082 (Proposed Standard), Internet Engineering Task Force, Oct. 2007. [Online]. Available: http://www.ietf.org/rfc/rfc5082.txt

17. S. Kent and R. Atkinson, "Security Architecture for the Internet Protocol," RFC 2401 (Proposed Standard), Internet Engineering Task Force, Nov. 1998, obsoleted by RFC 4301, updated by RFC 3168. [Online]. Available: http://www.ietf.org/rfc/rfc2401.txt

18. R. Rivest, "The MD5 Message-Digest Algorithm," RFC 1321 (Informational), Internet Engineering Task Force, Apr. 1992. [Online]. Available: http://www.ietf.org/rfc/rfc1321.txt

19. A. Heffernan, "Protection of BGP Sessions via the TCP MD5 Signature Option," RFC 2385 (Proposed Standard), Internet Engineering Task Force, Aug. 1998[Online]. Available: http://www.ietf.org/rfc/rfc2385.txt

20. M. Behringer, "BGP Session Security Requirements," Internet-Draft (Informational), Aug. 2007. [Online]. Available: http://tools.ietf.org/html/draft-behringer-bgp-session-sec-req-02

21. B. Christian and T. Tauber, "BGP Security Requirements," Internet-Draft (Informational), Nov. 2008. [Online]. Available: http://tools.ietf.org/html/draft-ietf-rpsec-bgpsecrec-10

22. B. Schneier, Applied Cryptography: Protocols, Algorithms, and Source Code in C, P. Sutherland, Ed. New York, NY, USA: John Wiley & Sons, Inc., 1995.

23. S. Murphy, "BGP Security Analysis," Nov. 2001. [Online]. Available: http://tools.ietf.org/html/draft-murphy-bgp-secr-04

24. D. Cooper, S. Santesson, S. Farrell, S. Boeyen, R. Housley, and W. Polk, "Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile," RFC 5280 (Proposed Standard), Internet Engineering Task Force, May 2008. [Online]. Available: http://www.ietf.org/rfc/rfc5280.txt

25. S. Kent, C. Lynn, and K. Seo, "Secure Border Gateway Protocol (SBGP),"Selected Areas in Communications, IEEE Journal on, vol. 18, no. 4, pp. 582–592, Apr 2000.

26. R. White, "Securing BGP Through Secure Origin BGP," The Internet Protocol Journal, vol. 6, no. 3, Sep 2003.

27. P. v. Oorschot, T. Wan, and E. Kranakis, "On interdomain routing security and pretty secure BGP (psBGP)," ACM Trans. Inf. Syst. Secur., vol. 10, no. 3, p. 11, 2007.

28. G. Goodell, W. Aiello, T. Griffin, J. Ioannidis, P. McDaniel, and A. Rubin, "Working Around BGP: An Incremental Approach to Improving Security and Accuracy of Interdomain Routing," in Proceedings of Internet Society Symposium on Network and Distributed System Security (NDSS 03), February 2003.

29. R. Perlman, "Network Layer Protocols with Byzantine Robustness," Tech. Rep., 1988. [Online]. Available: http://www.lcs.mit.edu/publications/specpub.php?id=997

30. B. Smith and J. Garcia-Luna-Aceves, "Securing the border gateway routing protocol," in Global Telecommunications Conference, 1996. GLOBECOM '96. 'Communications: The Key to Global Prosperity, Nov 1996, pp. 81–85

31. B. Smith and J. Garcia-Luna-Aceves, "Efficient security mechanisms for the border gateway routing protocol," Computer Communications, vol. 21, no. 3, pp. 203–210, March 1998.

32. T. Bates, R. Bush, T. Li, and Y. Rhekter, "DNS-based NLRI origin AS verification in BGP," Jul. 1998. [Online]. Available: http://tools.ietf.org/html/draft-bates-bgp4-nlri-orig-verif-00

33. Y.-C. Hu, A. Perrig, and D. Johnson, "Efficient Security Mechanisms for Routing Protocols," in Proceedings of Internet Society Symposium on Network and Distributed System Security (NDSS 03), February 2003

34. Y.-C. Hu, A. Perrig, and M. Sirbu, "SPV: secure path vector routing for securing BGP," in SIGCOMM '04: Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications. New York, NY, USA: ACM, 2004, pp. 179–192.

35. B. Raghavan, S. Panjwani, and A. Mityagin, "Analysis of the SPV secure routing protocol: weaknesses and lessons," SIGCOMM Comput. Commun. Rev., vol. 37, no. 2, pp. 29–38, 2007.

36. H. Chan, D. Dash, A. Perrig, and H. Zhang, "Modeling adoptability of secure BGP protocols," in SIGMETRICS '06/Performance '06: Proceedings of the joint international conference on Measurement and modeling of computer systems. New York, NY, USA: ACM, 2006, pp. 389–390.

37. K. Butler, T. R. Farley, P. McDaniel, and J. Rexford, "A Survey of BGP Security Issues and Solutions," Proceedings of the IEEE, vol. 98, no. 1, pp. 100–122, Jan. 2010.

38. G. Huston, M. Rossi and G. Armitage, "Securing BGP - A Literature Survey" IEEE communications surveys & tutorials, vol. 13, no. 2, second quarter 2011

39. Cisco Press, "How BGP works" 2008 [Online] Accessed March 2012-08-02 Available: http://www.cisco.com/en/US/tech/tk365/technologies_tech_note09186a00800c95bb.shtml#howbgpwork

40. M. Behringer, "BGP Session Security Requirements," Internet-Draft (Informational), Aug. 2007. [Online]. Available: http://tools.ietf.org/html/draft-behringer-bgp-session-sec-req-02

41. IETF, RFC 4272, "BGP Security Vulnerabilities Analysis", January 2006. http://www.ietf.org/rfc/rfc4272.txt

42. Docstoc, "Border gateway Protocol," [Online] Accessed May 2012 Available: http://www.docstoc.com/docs/6230899/Border_Gateway_Protocol

43. S. Deshpande, M. Thouttan and B. Sikdar "An Online Scheme for the Isolation of BGP Misconfiguration Errors," IEEE transactions on network and service management, vol. 5, no. 1, june 2008

44. P. McDaniel, W. Aiello, K. Butler, and J. Ioannidis, "Origin authentication in interdomain routing," Comput. Netw., vol. 50, no. 16, pp. 2953–2980, 2006.

45. J. Reason. *Human Error*. Cambridge University Press, 1990

46. L. Gao. On Inferring Autonomous System Relationships in the Internet. In IEEE Global Internet Symposium, Nov. 2000.

47. S. A. Misel. Wow, AS7007! NANOG mail archives. http://www.merit.edu/mail.archives/nanog/1997-04/msg00340.html.

48. Cisco Documentation, "Cisco Configuration Professional" [Online] Accessed June 2012. Available: http://www.cisco.com/en/US/products/ps9422/index.html

49. K. Butler, P. McDaniel, W. Aiello and J. Ioannidis, "Origin authentication in interdomain routing," Comput. Netw., vol. 50, no. 16, pp. 2953–2980, 2006.

50. IETF RFC 4301 – "Security Architecture for the Internet Protocol," December 2005, (obsoletes RFC 2401), also see companion RFCs 4302-4309, http://www.rfc-editor.org/rfc/rfc4301.txt

APPENDIX A – ACRONYMS

Selected acronyms used in this thesis are defined below

AS	Autonomous System
ASN	Autonomous System Number
BGP	Border Gateway Protocol
BGP-4	Border Gateway Protocol 4
CIDR	Classless Inter-domain Routing
DNS	Domain Name System
EBGP	Exterior Border Gateway Protocol
EGP	Exterior Gateway Protocol
IGP	Interior Gateway Protocol(e.g., iBGP, OSPF, RIP)
OSPF	Open Shortest Path First
EIGRP	Enhanced Interior Gateway Routing Protocol
GTSM	Generalized TTL Security Mechanism
IRV	Inter-Domain Route Validation
IBGP	Internal Border Gateway Protocol
TCP	Transmission Control Protocol
IETF	Internet Engineering Task Force
IP	Internet Protocol
IPsec	Internet Protocol Security
Ipv4	Internet Protocol version 4
Ipv6	Internet Protocol version 6
ISP	Internet Service Provider
NAP	Network Access Point
LAN	Local Area Network
MED	Multi-Exit Discriminator
ACL	Access Control LIst
RFC	Request For Comment
MD5	Message Digest Algorithm – Version 5
DNS	Domain Name System
SBGP	Secure Border Gateway Protocol
psBGP	Pretty Secure Border Gateway Protocol
soBGP	Secure Origin Border Gateway Protocol
MAC	Message Authentication Code
SPV	Secure Path Vector routing
NLRI	Network Layer Reachability Information
GNS 3	Graphical Network Simulator 3
NS2	Network Simulator 2
OPNET	Optimized Network Engineering Tools
VoIP	Voice Over Internet Protocol

QoS	Quality of Service
MPLS	Multi Protocol Label Switching