Understanding and Detecting BGP Instabilities

Roma Tre University

Advisor: Prof. G. Di Battista
PhD Candidate: Luca Cittadini
BGP glues the Internet

- Internet is a huge interconnection
  - a network of networks
- BGP is the protocol that makes interconnections possible
  - disseminates routing information among heterogeneously administered networks (ASes)
  - makes networks aware of each other
- BGP is extremely hard to upgrade/replace
  - need to deploy a new protocol worldwide
  - huge legacy installation base (30k ASes)
BGP instabilities

- BGP is designed to
  - fulfill the classic goals of routing protocols...
    - build “optimal” routing tables, avoid loops, etc.
  - support detailed routing policies
    - administrators must have fine-grained knobs to control how traffic enters/exits their network

- policy conflicts can create instabilities
  - aka oscillations
  - transient (permanent) situations where routers are unable to reach a fixed set of routing choices
motivation

• BGP instabilities are harmful
  • can generate a very large (infinite) amount of messages
  • can delay convergence
• instabilities are hard to fix
  • the “cure”, i.e., rate limiting routing updates, is worse than the disease
  • motivates efforts to prevent oscillations
a model for BGP

- we choose SPVP [GriffinShepherdWilfong99]
  - an undirected graph represents BGP peerings
  - a single destination prefix is originated by node 0
- each node is assigned a set of permitted paths to reach 0 (filtering component)
- paths at \( \nu \) are sorted according to preference (ranking component)
SPVP – dynamic model

- original version
  - node 0 advertises its presence to its neighbors
  - each node
    - collects paths from neighbors
    - applies filters to received paths
    - selects the highest ranked available path
    - updates its neighbors

- many simplified variants proposed in literature
  - nodes cannot talk simultaneously
  - nodes send/receive paths at each clock tick
  - paths are not stored locally

- result: simplified variants are unable to capture all BGP oscillations
BGP research - coordinates

network management

feasibility of finding solutions

stability vs autonomy and expressiveness
agenda

- sufficient AND necessary condition for stability
- impact of BGP attribute manipulation
- static analysis of BGP configurations
safety under filtering

- A network is safe under filtering (SUF) if it is guaranteed to converge to a stable routing even if arbitrary route filters are applied.
- A network is robust if it is guaranteed to converge to a stable routing even under arbitrary combinations of link failures.

results
- Robustness does not imply SUF.
  - Route filters can be more harmful than cable cuts.
- Characterization for SUF.
  - Does not depend on dynamics (hence, can be checked statically).
wheels

- A **Dispute Wheel** is a cyclic structure of preferences:
  - the structure is made of pivot nodes
  - each pivot has a direct route
  - each pivot has a route via its successor
  - each pivot prefers the route via its successor to the direct route

- No Dispute Wheel $\implies$ SUF [GriffinShepherdWilfong99,02]
rings

- A Dispute Ring is a DW such that each node appears only once in the wheel
- SUF => No Dispute Ring
  [FeamsterJohariBalakrishnan05]
- Intuition
  - meet in the middle to characterize SUF

too complex  right!  too simple
A Dispute Reel (DR) is a particular kind of DW and a generalization of a Dispute Ring.

A DR is a DW such that
1. Pivot vertices appear in exactly three paths
2. Spoke and rim paths do not intersect
3. Spoke paths form a tree
   • only intersections among rim paths are allowed

A DW that does not satisfy these conditions does not pose stability problems
the “big picture”

NO DISPUTE REEL

SAFE

NO DW

HAS A
STABLE
STATE

SUF

ROBUST

NO Dispute Ring

Filthy Gadget
characterization of SUF

network management

feasibility of finding solutions

stability vs autonomy and expressiveness
agenda

- sufficient AND necessary condition for stability
- impact of BGP attribute manipulation
- static analysis of BGP configurations
iBGP attribute manipulation

- internal BGP distributes routes within an AS
- vendors do not recommend applying policies to routes learned via iBGP
  - yet, there are traffic balancing reasons to do so
  - e.g., when you want multiple routes to survive the BGP decision process up to local tie breakings
- consequences are poorly understood
iBGP - results

- measurement methodology
  - exploits the simultaneous availability of uncomparable paths at the same AS
  - iBGP attribute manipulation happens in the Internet

- theoretical analysis
  - arbitrary manipulation can create oscillations which are not possible otherwise

- configuration guidelines
  - match reasonable traffic engineering requirements while ensuring that convergence is preserved
  - the iBGP-equivalent of Gao-Rexford guidelines
impact of iBGP attr. manip.

network management

feasibility of finding solutions

stability vs autonomy and expressiveness
agenda

- sufficient AND necessary condition for stability
- impact of BGP attribute manipulation
- static analysis of BGP configurations
the greedy+ algorithm

- **intuition:**
  - Some paths, e.g. “O”, are guaranteed
  - Paths that are worse ranked than guaranteed paths will not be selected

- **algorithm:**
  - iteratively grow a set of Stable nodes
  - pin guaranteed routes
  - purge less preferred and unfeasible paths
an automatic BGP convergence checker

- Collect and parse configs
- Translate to SPVP
- Run Greedy+

- Easy, e.g. SNMP, JunXML,...
- May take exponential time
- Polynomial time
translation to SPVP

- idea: prune unnecessary paths
  - simulate announcement propagation to generate paths
  - exploit Greedy+ pruning steps to make the path generation process smarter
- some nodes will be stabilized during the generation
  - We generate only one path for "early stabilized" nodes
- some paths will be less preferred than stable paths
  - We do not generate them
results

- **theoretically**
  - A deterministic P-time greedy heuristic to check whether a configuration potentially admits an oscillation
    - No false-negatives: never misreports a network as stable

- **practically**
  - An efficient way to map configurations (even for Internet-scale networks) to the abstract SPVP model
  - An efficient way to check the SPVP network for potential oscillations
  - In the worst case, 0.3 sec to check the stability of a large iBGP network for a given destination network
BGP Analyzer - Architecture

BGP2SPVP translator

Node 1

Router Configurations

Router RIBs

IGP Link Weights

SPVP instance

SPP Graph

Permitted Paths

Ranking Functions

Configuration Parser

iBGP Topology

Dissemination Phase

Ranking Phase

RIB Parser

IGP Parser
static analysis of BGP policies

network management

feasibility of finding solutions

stability vs autonomy and expressiveness
other research topics

- clean-slate routing architectures
  - joint work with Anja Feldmann’s group @ TU Berlin
- IPv4-IPv6 transition and coexistence
  - joint work with Olaf Maennel (Univ. of Loughborough), Randy Bush (IIJ), et al.
- IPv4 address space usage
  - joint work with Wolfgang Muhlbauer (ETH Zurich) and Steve Uhlig (TU Berlin)
main achievements

- unrestricted local policies are intrinsically incompatible with guaranteed convergence
  - we must sacrifice expressiveness to preserve filtering autonomy and prevent oscillations
  - even in iBGP, increased expressiveness implies increased risk of oscillations
- stability can be analyzed statically
  - polynomial-time algorithm on SPVP instances
    - no false positives
  - efficient translation from Internet-like topologies to SPVP instances
thank you!

- any questions?
Backup Slides

(the gory details)
classes of SPVP instances
greedy+ - an example (1)
greedy+ – an example (2)

Inconsistent: extends an unavailable route
greedy+ - an example (3)
policy checker - performance

Naive Greedy Greedy+ (min-max)

4M paths (out of memory)

Number of generated paths (median)

Topologies (degree threshold)

degree >1000
7 ASes, 21 links

degree >126K
ASes, 60K links

degree >1
85
26K ASes, 60K links
performance (iBGP)
L. Cittadini, W. Muhlbauer, S. Uhling. Evolution of Internet Address Space Deaggregation: Myths and Reality. IEEE Journal on Selected Areas in Communications (JSAC) Special Issue on Internet Routing Scalability, 2010
conference publications


conference publications

conference publications

internet draft