How DNS and BGP work to serve your Internet

Contents

[**Border Gateway Protocol** 2](#_Toc185580534)

[BGP, Autonomous Systems, and Facebook Outage 5](#_Toc185580535)

[Facebook Outage Analysis: 7](#_Toc185580536)

## **Border Gateway Protocol**

**Introduction to BGP:**

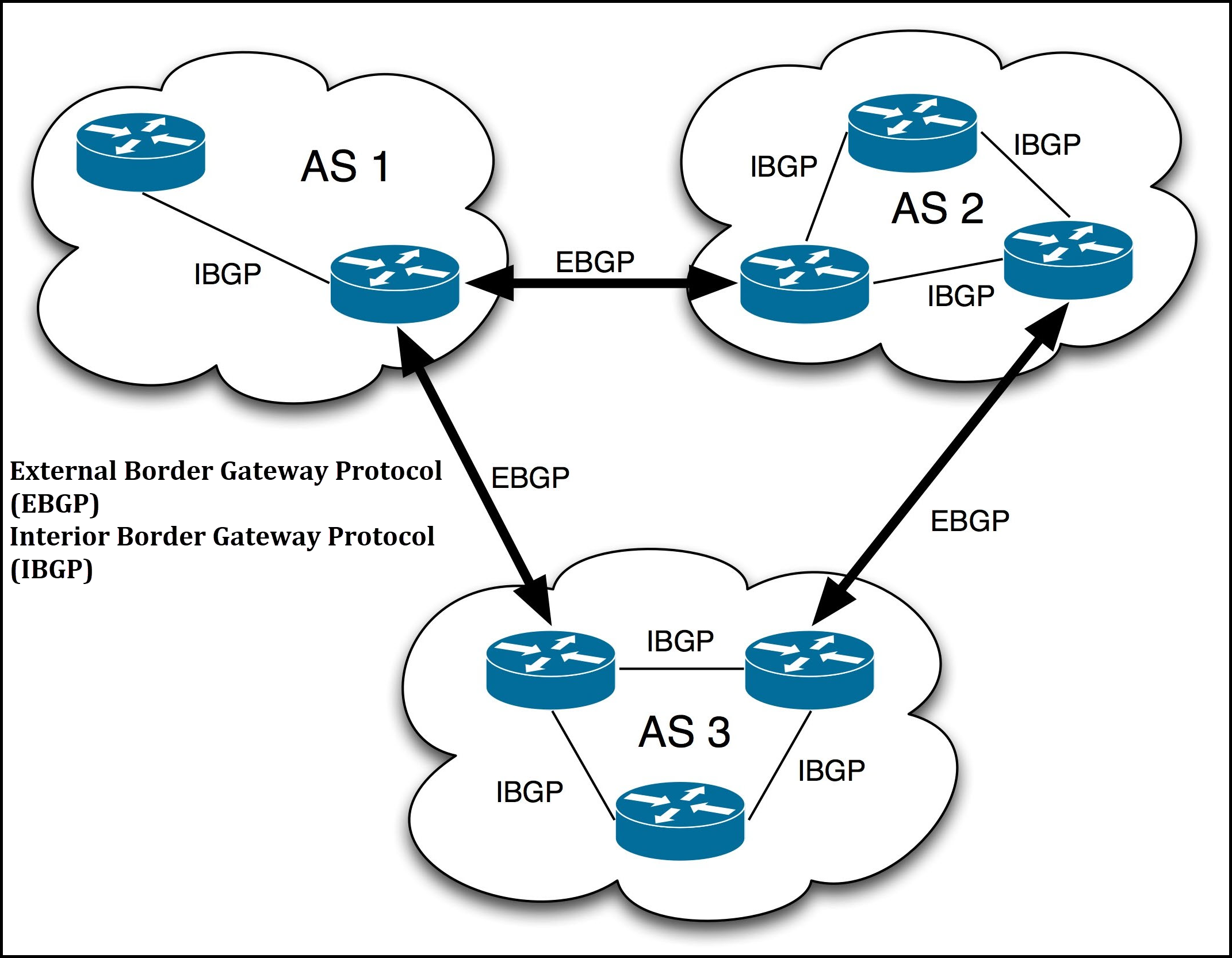
* BGP stands for **Border Gateway Protocol**.
* BGP is a protocol that facilitates the routing of data packets between different networks on the Internet, forming a "network of networks."
* Each network in BGP terminology is known as an **Autonomous System (AS)**.

**Analogies to Understand BGP:**

1. **Human Communication Analogy**:
   * If you know the address of a person, you can send a package via a postal service.
   * The postal service determines the best route to deliver the package based on various factors like speed, cost, and commitments.
2. **Courier Network Analogy**:
   * Packages move through tiered systems:
     + From a small town to a nearby tier-1 city or warehouse.
     + Then to the tier-1 city closest to the destination.
     + Finally, to the recipient's smaller town.
   * The Internet works similarly, with data packets traveling through interconnected networks.
3. **Network of Networks**:
   * Autonomous Systems (AS) are like cities or hubs in a courier network, interconnected to form the global Internet.

**Key Concepts in BGP:**

1. **Autonomous Systems (AS):**
   * An AS is a single network or a group of networks under a common administrative domain.
   * Examples: ISPs (e.g., Jio, Airtel), universities (e.g., Stanford, USC), and tech companies (e.g., Google, Facebook).
2. **AS Number (ASN):**
   * Each AS is uniquely identified by an ASN.
   * ASNs are assigned by organizations like IANA.
     + The Internet Assigned Numbers Authority (IANA) is a standards organization that oversees global IP address allocation, autonomous system number allocation, root zone management in the Domain Name System (DNS), media types, and other Internet Protocol–related symbols and Internet numbers.
   * Examples:
     + Facebook owns certain ASNs.
     + Google has multiple ASNs for its networks.



1. **Factors for Routing Decisions:**
   * Speed (fastest path).
   * Cost (economic considerations for data transfer).
   * Priority (e.g., commitments for guaranteed delivery times).
   * Security and other policies.
2. **BGP Characteristics:**
   * BGP helps decide the best path between autonomous systems.
   * Path selection considers factors like costs associated with data transfer and time efficiency.

**BGP Operations:**

1. **Peering Sessions:**
   * Routers within autonomous systems share information about their network and paths with their neighbours.
   * This information sharing is called **advertisement**.
2. **Path Discovery:**
   * Initially, each router knows only about its neighbours.
   * During peering sessions, routers exchange knowledge about reachable networks.
   * Over multiple iterations, every router learns about the paths to all networks in the system.

**Important Terms:**

1. **Autonomous System (AS):**
   * A network or collection of networks managed as a single entity.
2. **ASN:**
   * Unique identifier for each AS.
3. **Peering:**
   * The process where routers exchange routing information with neighbours.
4. **Advertisement:**
   * Sharing of routing information between autonomous systems.

**Real-World Implications:**

1. **Costs and Agreements:**
   * ISPs and companies negotiate the cost of routing traffic through their networks.
   * Example: Google might charge Jio for using its AS to route data.
2. **Internet Infrastructure:**
   * Data doesn’t travel randomly; it follows paths defined by BGP based on agreements, policies, and optimizations.

**Key Takeaways:**

* BGP ensures data packets find optimal paths across a "network of networks."
* Autonomous Systems (AS) form the backbone of the Internet, with each AS managed by ISPs, universities, or companies.
* BGP decisions are based on speed, cost, and policies.
* The protocol relies on peering sessions and advertisements for path discovery and selection.

## BGP, Autonomous Systems, and Facebook Outage

**Key Concepts:**

**Autonomous Systems (AS):**

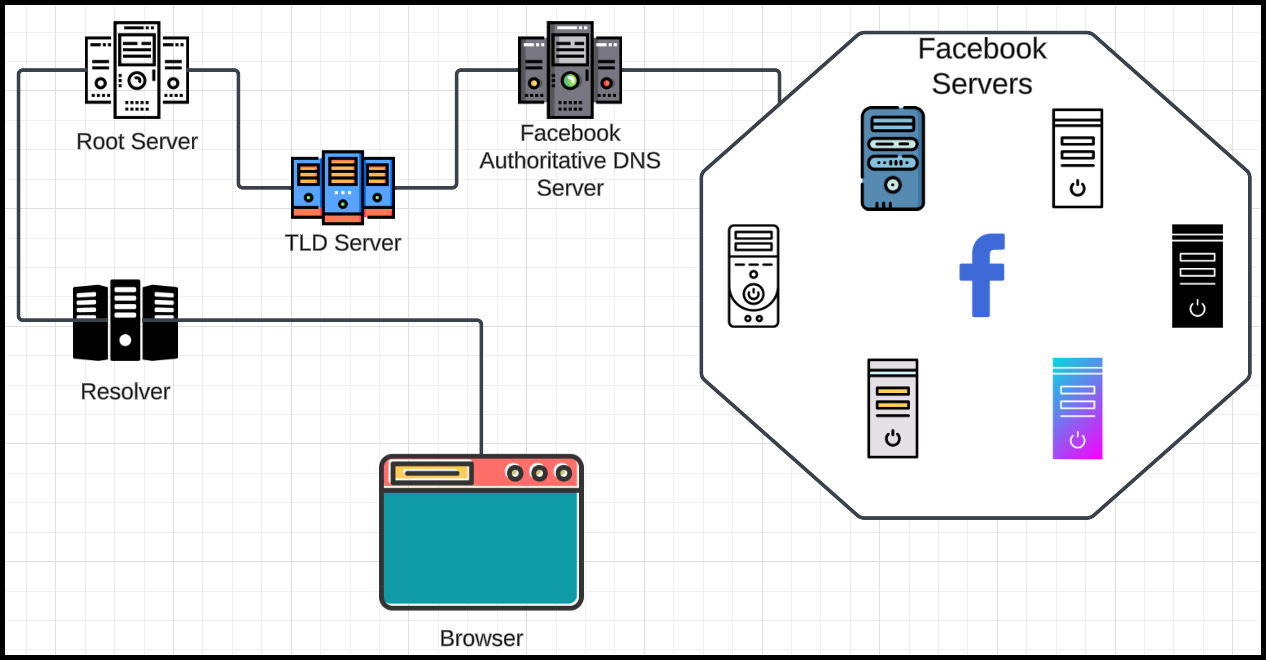
* An **Autonomous System (AS)** is a collection of connected IP routing prefixes under the control of a single organization.
* Each AS has information about its internal network as well as paths to other networks.
* AS use the **Border Gateway Protocol (BGP)** to communicate and advertise routes to other AS.
* BGP advertisements share information like the next hop, network capacity, and bandwidth to facilitate packet routing.

**Role of DNS and BGP:**

* **DNS (Domain Name System):** Translates domain names (e.g., facebook.com) into IP addresses.
  + Root servers, TLD servers, and authoritative servers work in sequence to resolve domain names.
  + Example: Resolver → Root DNS → TLD DNS → Authoritative DNS → IP address.
* **BGP:** Determines the best path to reach the destination IP.
  + BGP advertisements provide routing information about available paths.
  + If a route is not advertised, the path cannot be discovered, and communication fails.

**How Internet Routing Works:**

1. **DNS Resolution:**
   * User’s browser queries a DNS resolver to resolve a domain name.
   * Resolver communicates with root, TLD, and authoritative DNS servers to get the IP address.
2. **Path Determination:**
   * After obtaining the IP, packets are routed using BGP to determine the best path between autonomous systems.



## Facebook Outage Analysis:

**Cause:**

* Facebook's DNS servers stopped advertising routes to their autonomous systems via BGP.
* A configuration change disconnected Facebook’s DNS servers from its internal infrastructure, causing DNS queries to fail.

**Sequence of Events:**

1. **Initial Query:**
   * Users queried DNS resolvers for facebook.com.
   * Resolver attempted to reach Facebook’s authoritative DNS servers via BGP.
2. **BGP Failure:**
   * Facebook’s DNS servers stopped advertising their presence to BGP due to internal misconfiguration.
   * Without BGP advertisements, no path existed to reach Facebook’s DNS servers.
3. **Cascading Effects:**
   * Without DNS resolution, users could not obtain IP addresses for Facebook services.
   * Even internal debugging tools hosted on Facebook’s domains became inaccessible to engineers.

**Why Facebook Designed Its System This Way:**

* **Autonomous System Design:**
  + Facebook’s DNS servers are in separate autonomous systems for reliability.
  + If all servers in a particular AS become unreachable, the DNS stops advertising that AS to avoid routing users to unavailable servers.
* **Load Balancing via DNS:**
  + Authoritative DNS servers randomly return IP addresses of available servers for load distribution.

**Issues with Configuration Update:**

* The misconfiguration disrupted communication between DNS servers and internal infrastructure.
* The BGP advertisements were withdrawn, causing the network to become unreachable globally.
* Engineers faced additional delays because their debugging tools were also inaccessible.

**Lessons Learned:**

1. **Incremental Updates:**
   * Configuration changes should be rolled out regionally or incrementally to detect issues early.
2. **Separate Debugging Systems:**
   * Critical debugging tools should be hosted outside the primary network to ensure accessibility during outages.
3. **Redundancy and Testing:**
   * Ensure redundant paths and rigorous testing before applying global configurations.

**Summary:**

The Facebook outage was caused by a misconfigured update that disrupted DNS connectivity and BGP advertisements. This resulted in global inaccessibility to Facebook’s services. Proper planning, redundancy, and separation of critical systems are essential to avoid such issues.