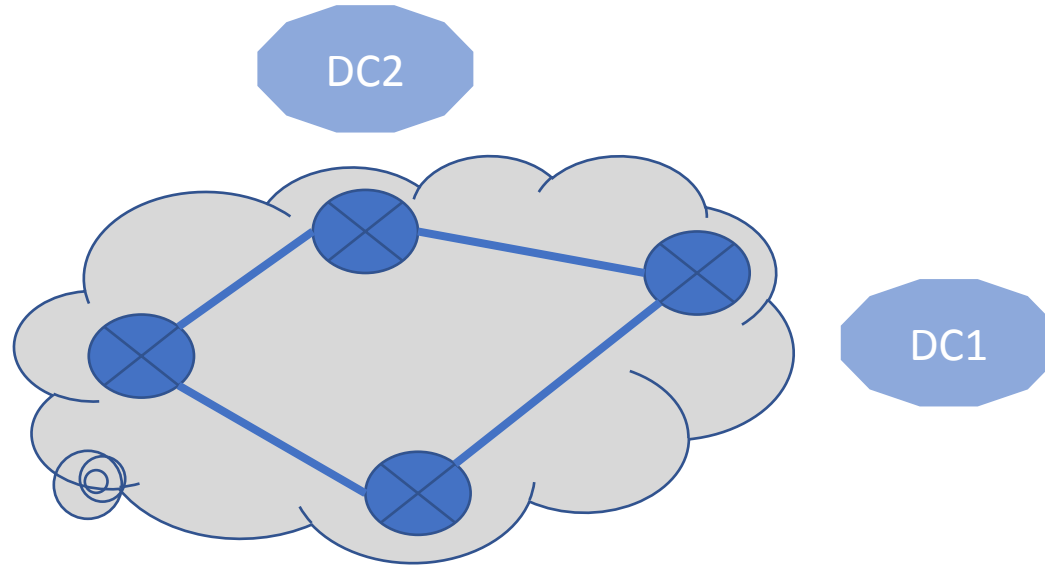


CSE 550: *Systems for all*

Au 2021

Ratul Mahajan

FB outage: Murder, suicide, and obstruction



FB outage: Murder, suicide, and obstruction

Normal operation

FB has a global backbone that connects its DCs and a distributed DNS infra.

DNS servers measure “distance” to different DCs.

DNS offers “close” DC prefixes to users and withdraws (from BGP) for unreachable DCs

Murder

An engineer or a script sends a bad command to backbone routers.

Audit tool fails to detect the bad command.

All DCs are disconnected from the backbone.

Suicide

DNS can no longer any DC (since they are all disconnected).

DNS withdraws many prefixes from BGP.

The prefixes cover the DNS infra as well, so DNS makes itself unreachable as well.

Obstruction

Service restoration requires manual intervention (since nothing is reachable).

Physical access requires authorization that turn depends on the same DNS.

Takes multiple hours to override systems and gain access to the equipment.

Lessons?

Impact of independent parties in BGP

1. Security
2. Performance
3. Convergence

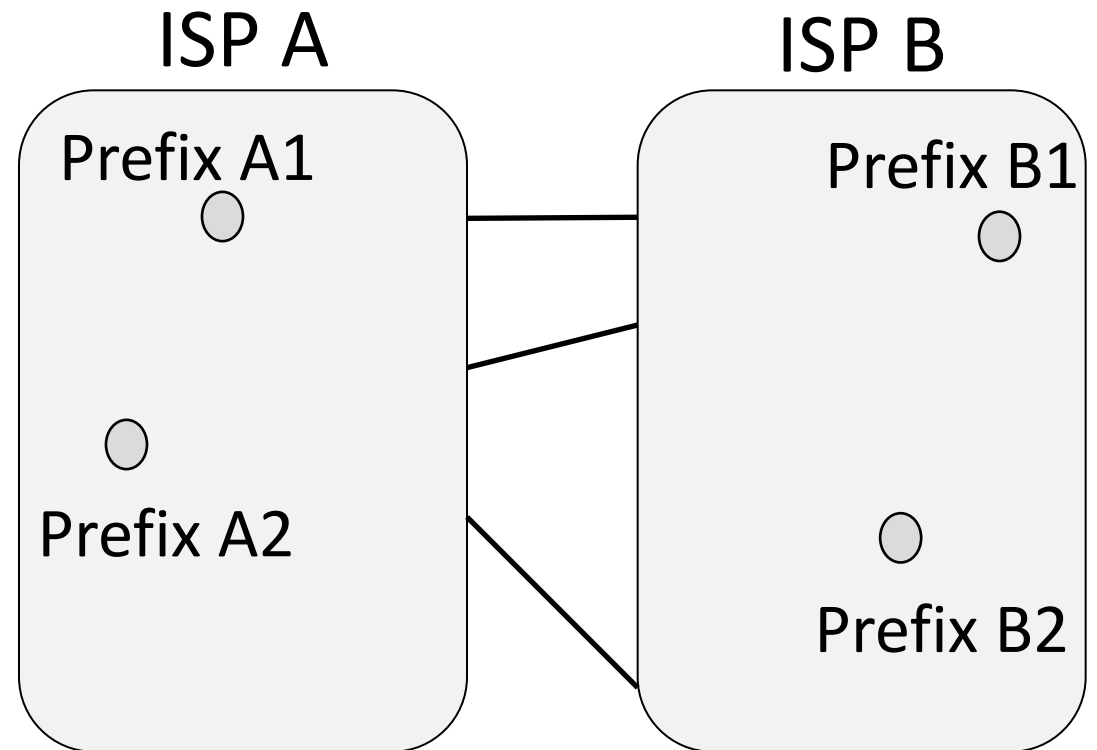
BGP performance

Each party selects routes to suit its own interests

- e.g, shortest path in ISP

What path will be chosen for $A2 \rightarrow B1$ and $B1 \rightarrow A2$?

- What is the best path?

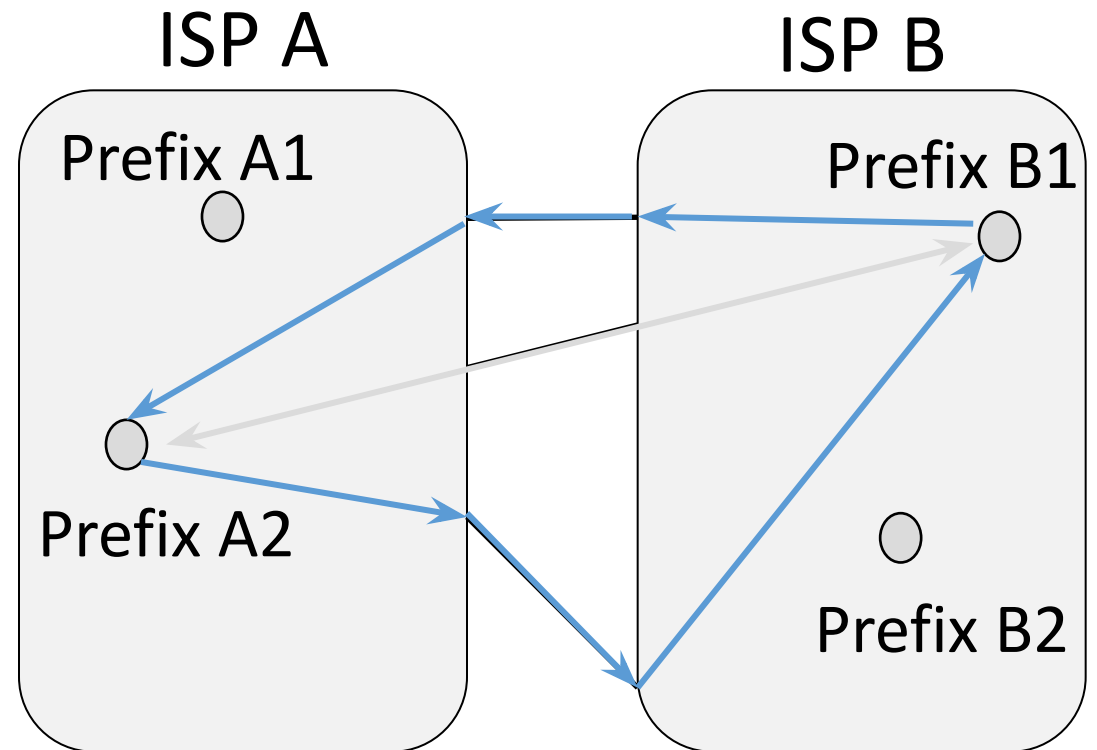


BGP performance

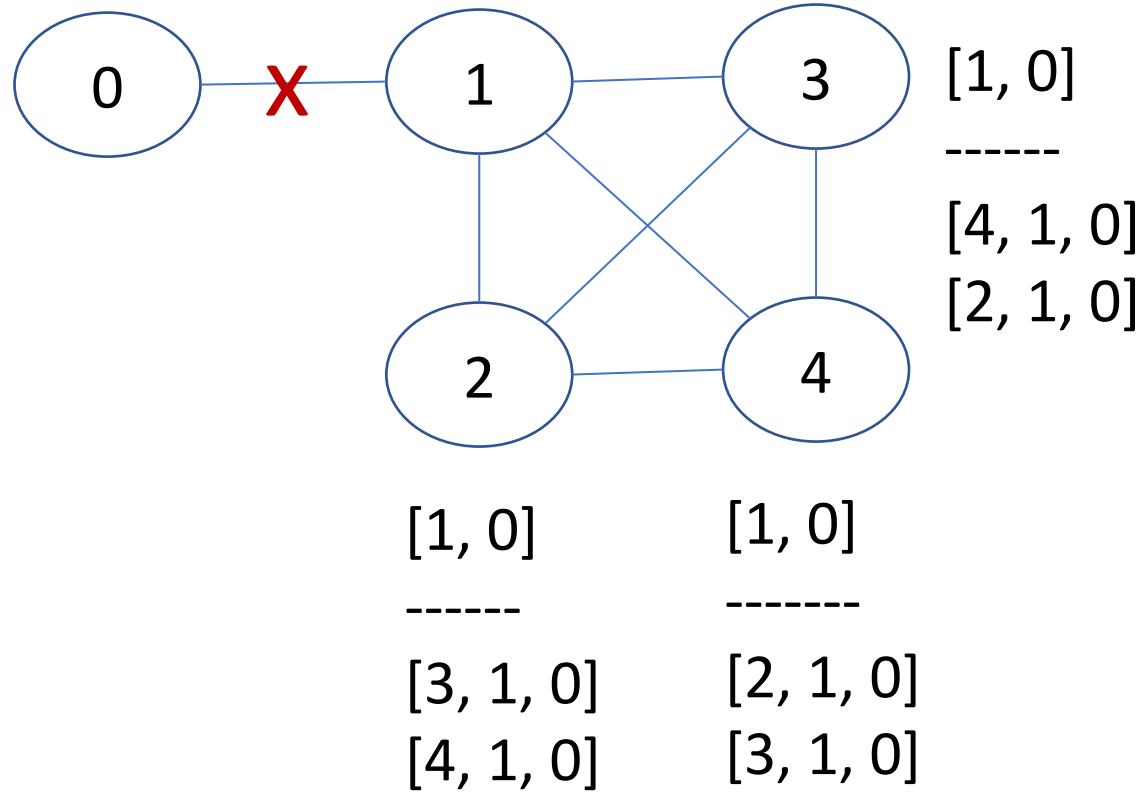
Selected paths are longer than overall shortest path

- And asymmetric too!

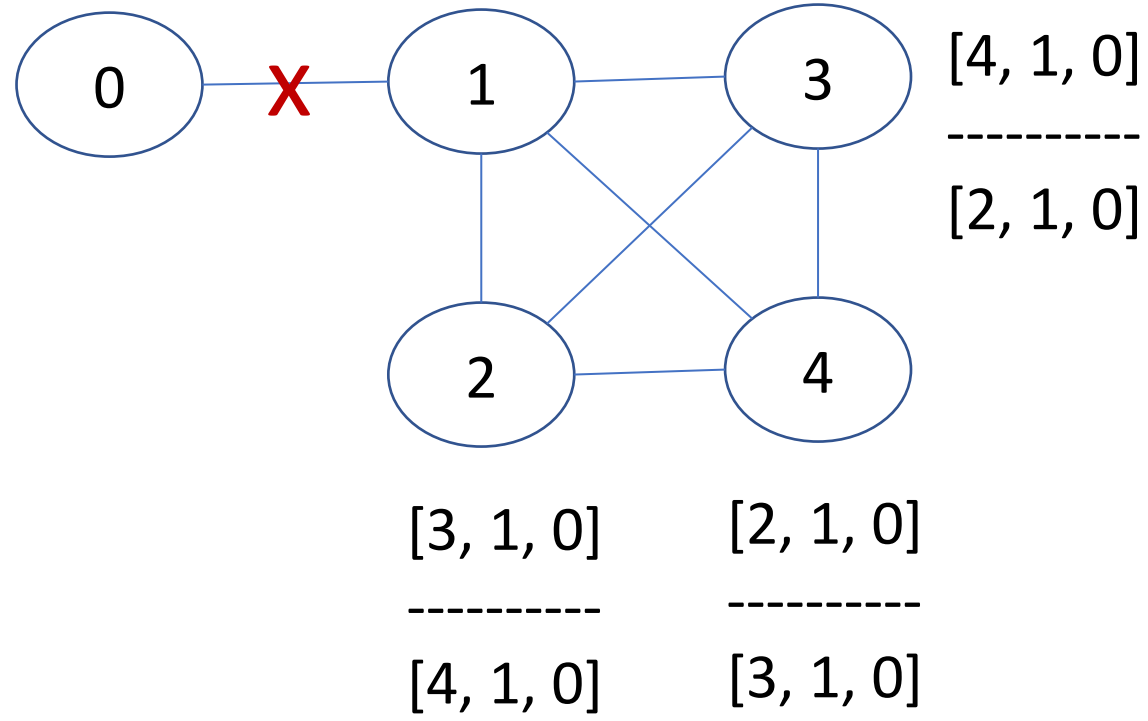
Consequence of independent goals and decisions (not hierarchy and info hiding)



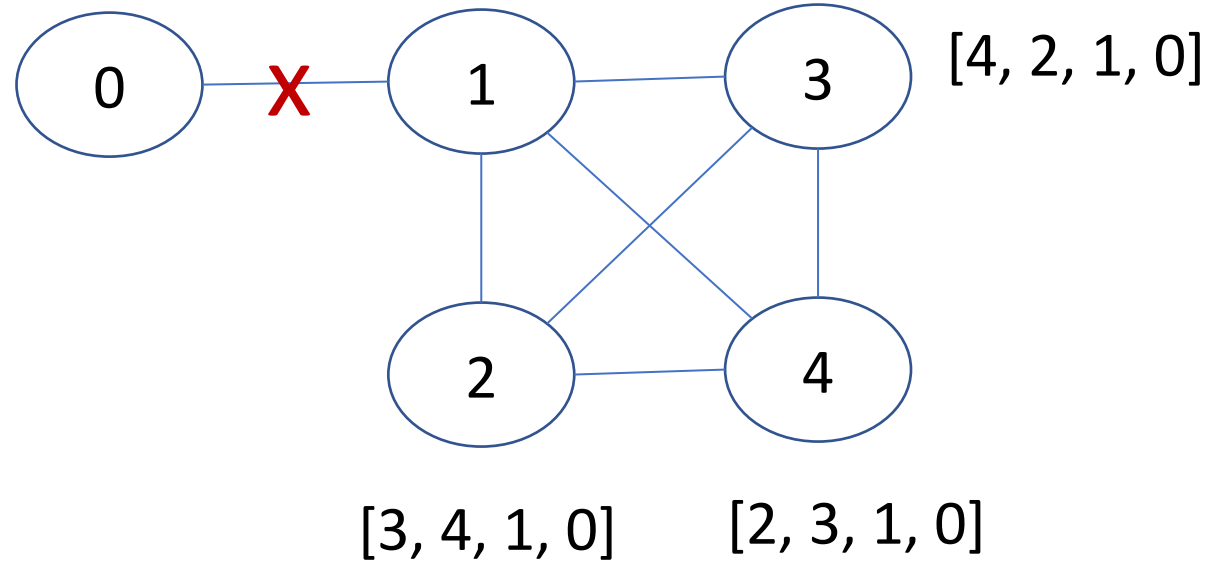
BGP slow convergence



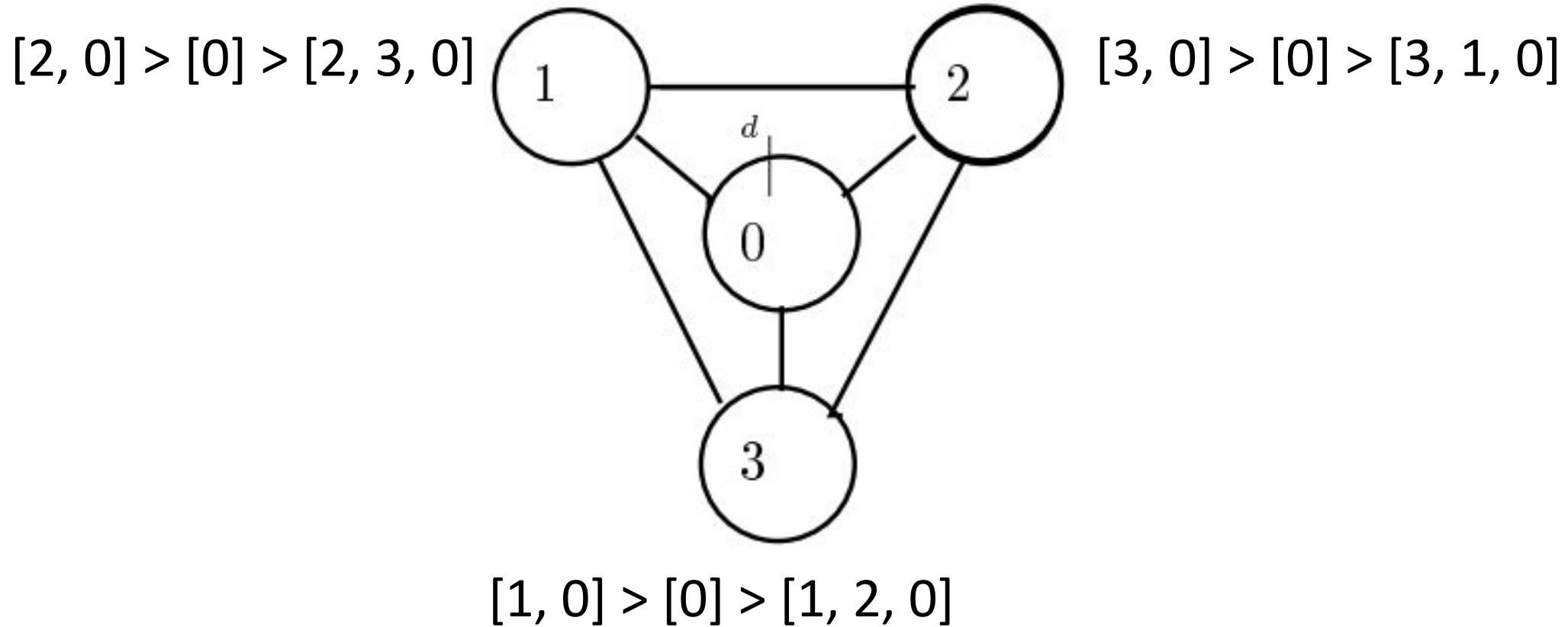
BGP slow convergence



BGP slow convergence



BGP “bad gadget”: Non-convergence



Over to Alice and Meng-Li