The Role of Routing Policies in the Internet: Stability, Security, and Load-Balancing

candidate: Marco Chiesa
advisor: prof. Giuseppe Di Battista

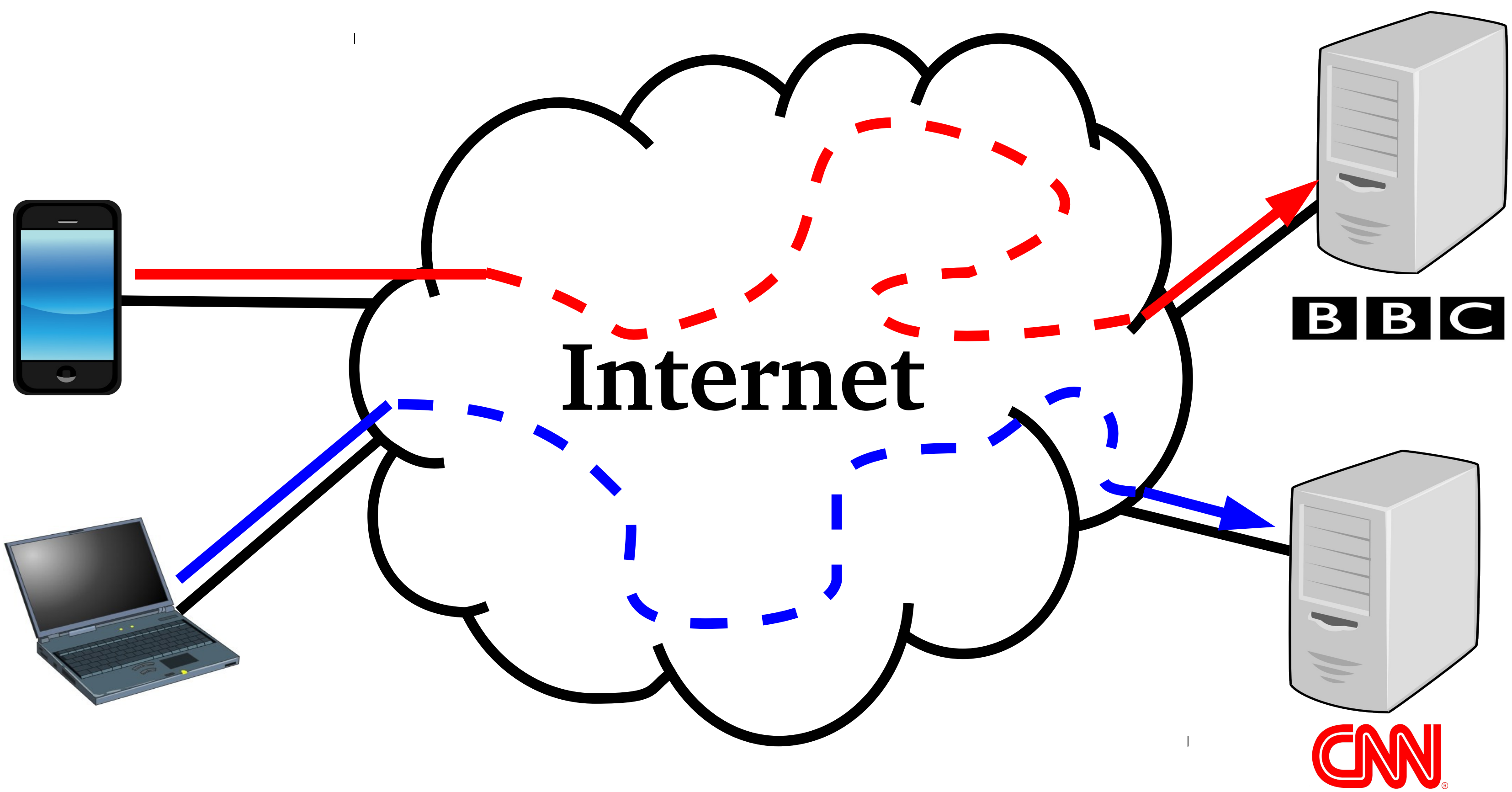
XXVI Ciclo

agenda

- Internet routing
- three challenges
- I. stability
- II. security
- III. load-balancing
- conclusions

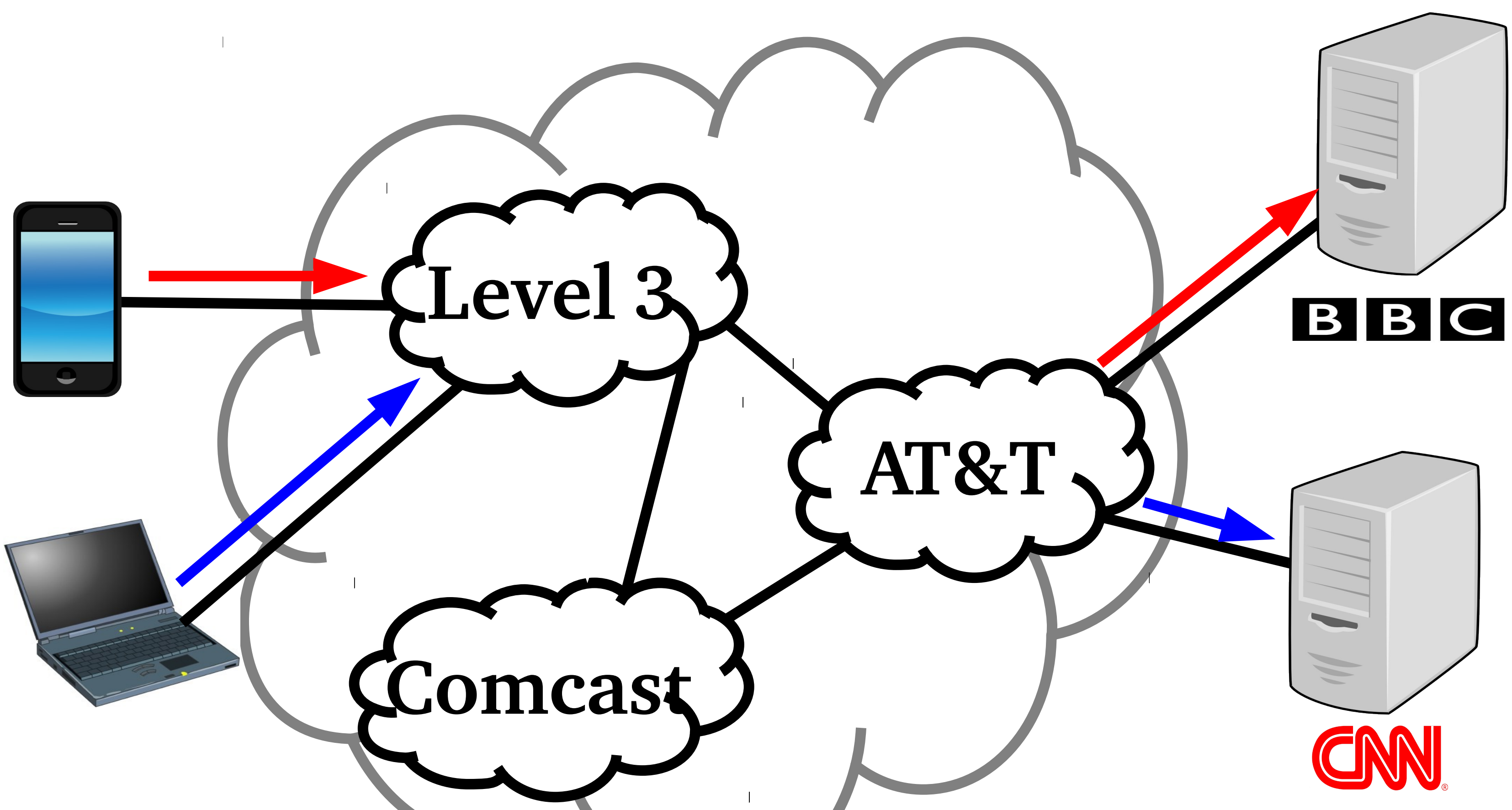
Internet routing

goal: provide connectivity among Internet devices



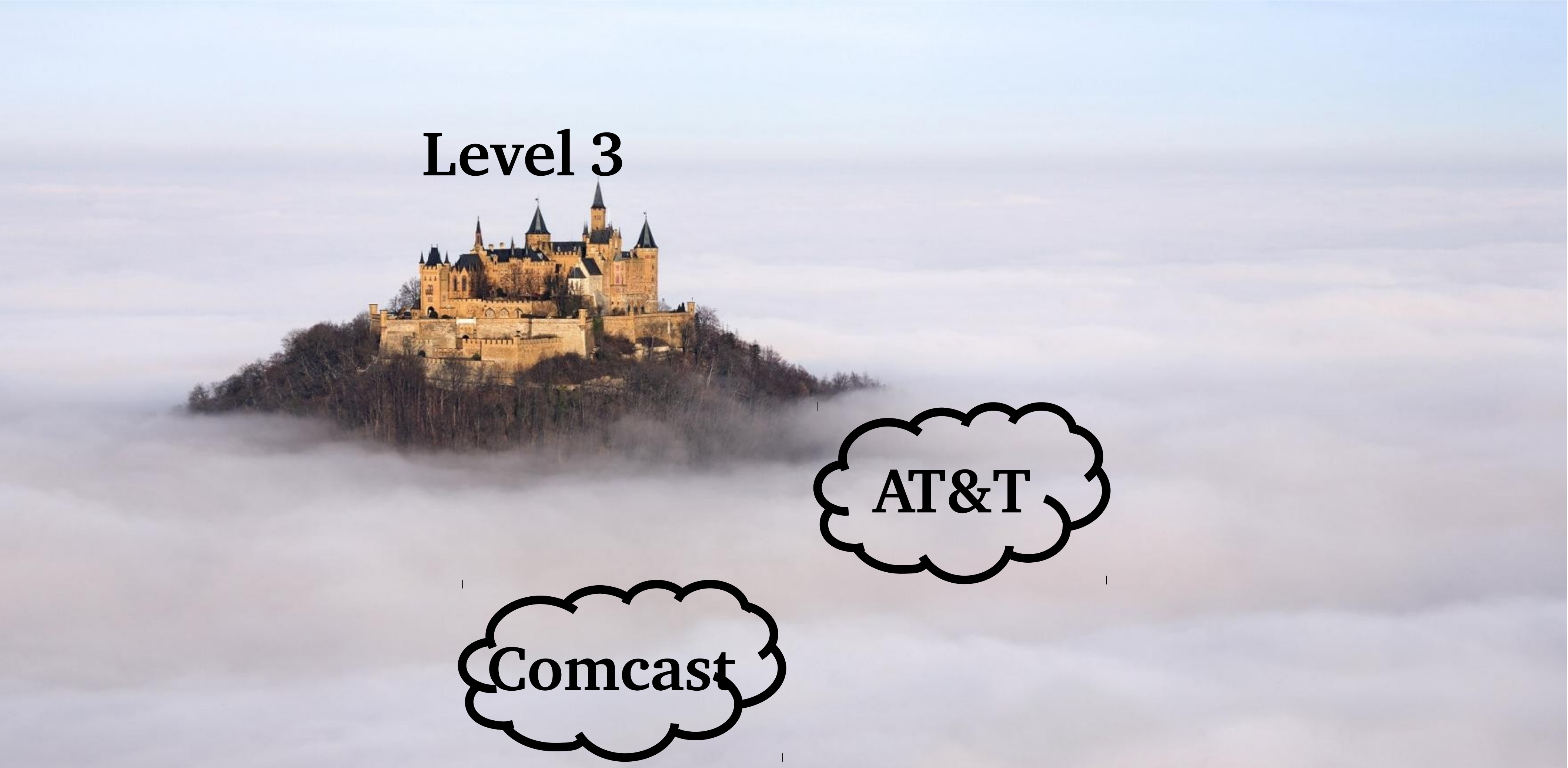
interdomain routing

- autonomous entities (Internet Service Providers, ISP)
- distributed (Border Gateway Protocol, BGP)



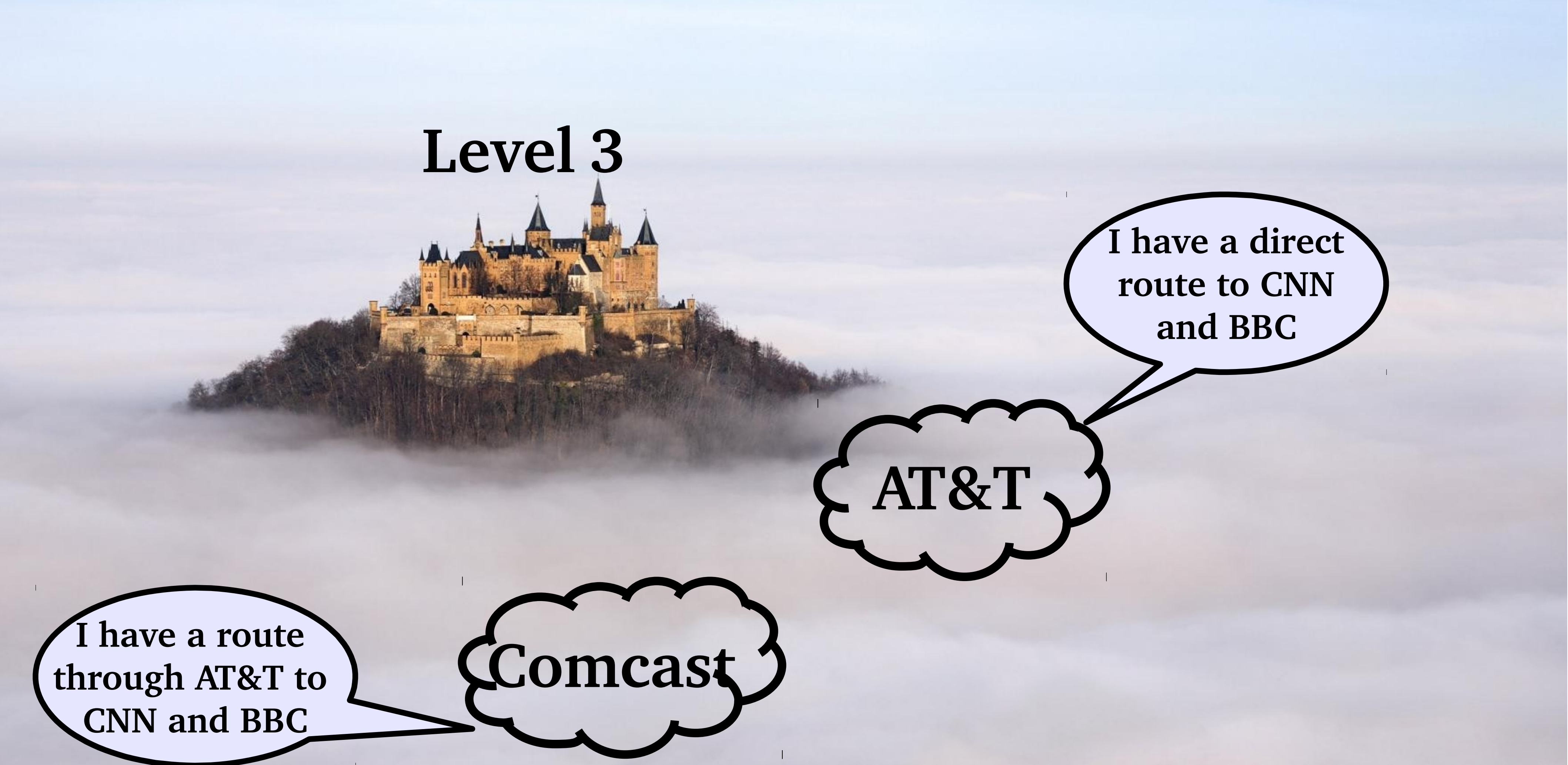
interdomain routing: Level 3 perspective

- autonomous entities (Internet Service Providers, ISP)
- distributed (Border Gateway Protocol, BGP)



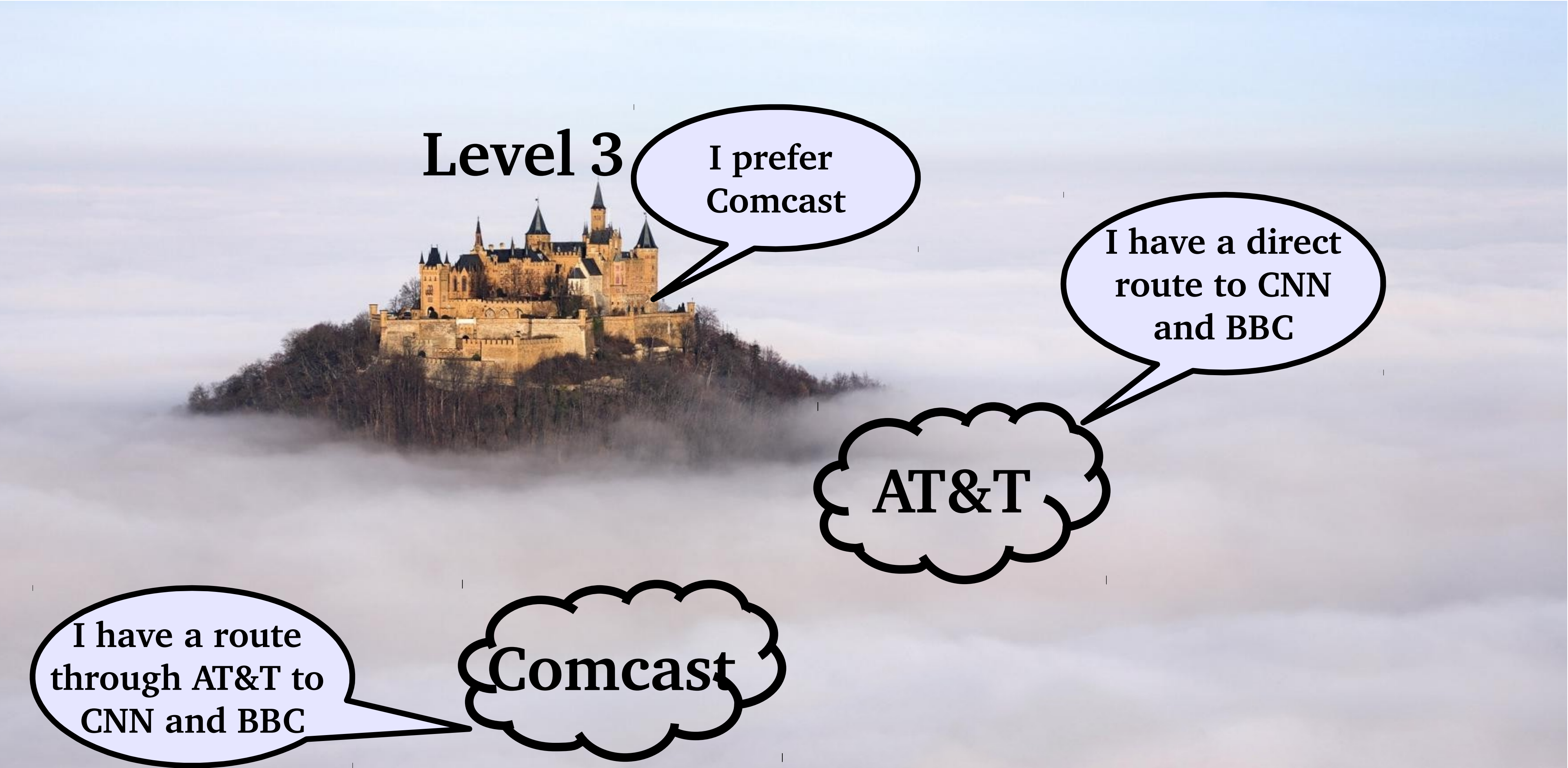
interdomain routing: Level 3 perspective

- autonomous entities (Internet Service Providers, ISP)
- distributed (Border Gateway Protocol, BGP)



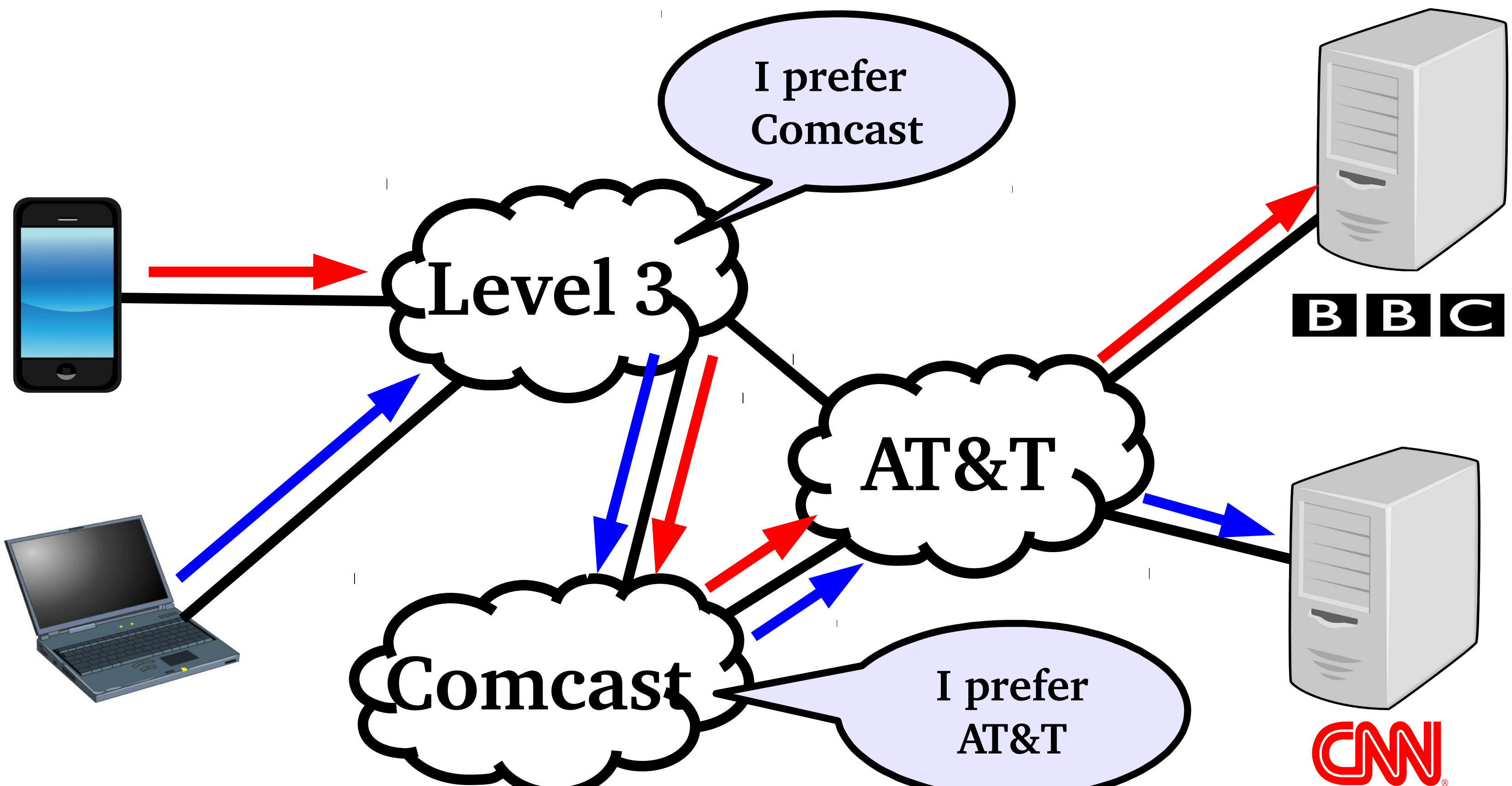
interdomain routing:
Level 3 perspective

- autonomous entities (Internet Service Providers, ISP)
- distributed (Border Gateway Protocol, BGP)



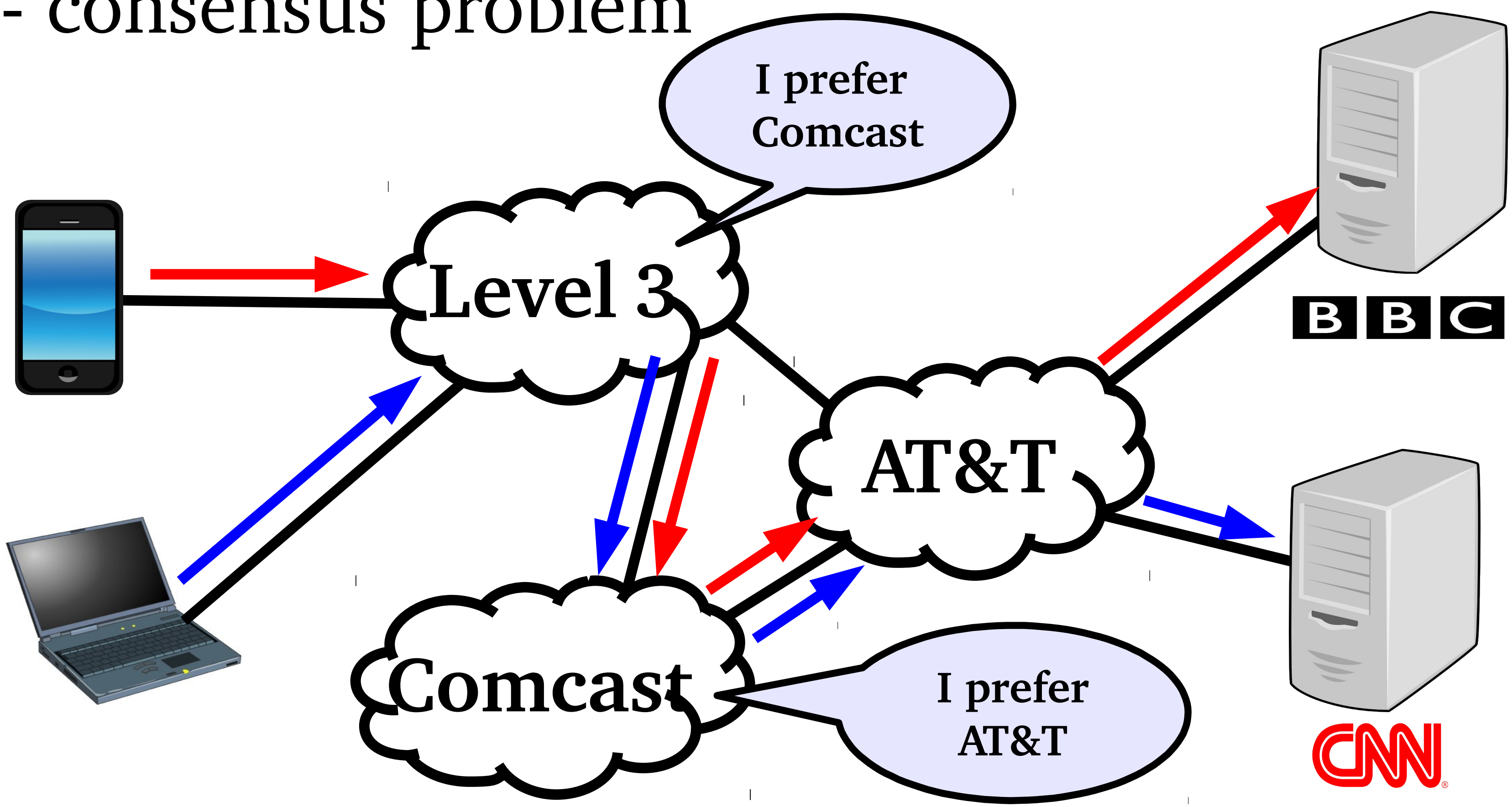
interdomain routing: routing policies

- autonomous entities (Internet Service Providers, ISP)
- distributed (Border Gateway Protocol, BGP)
- economic relationships ($$$)



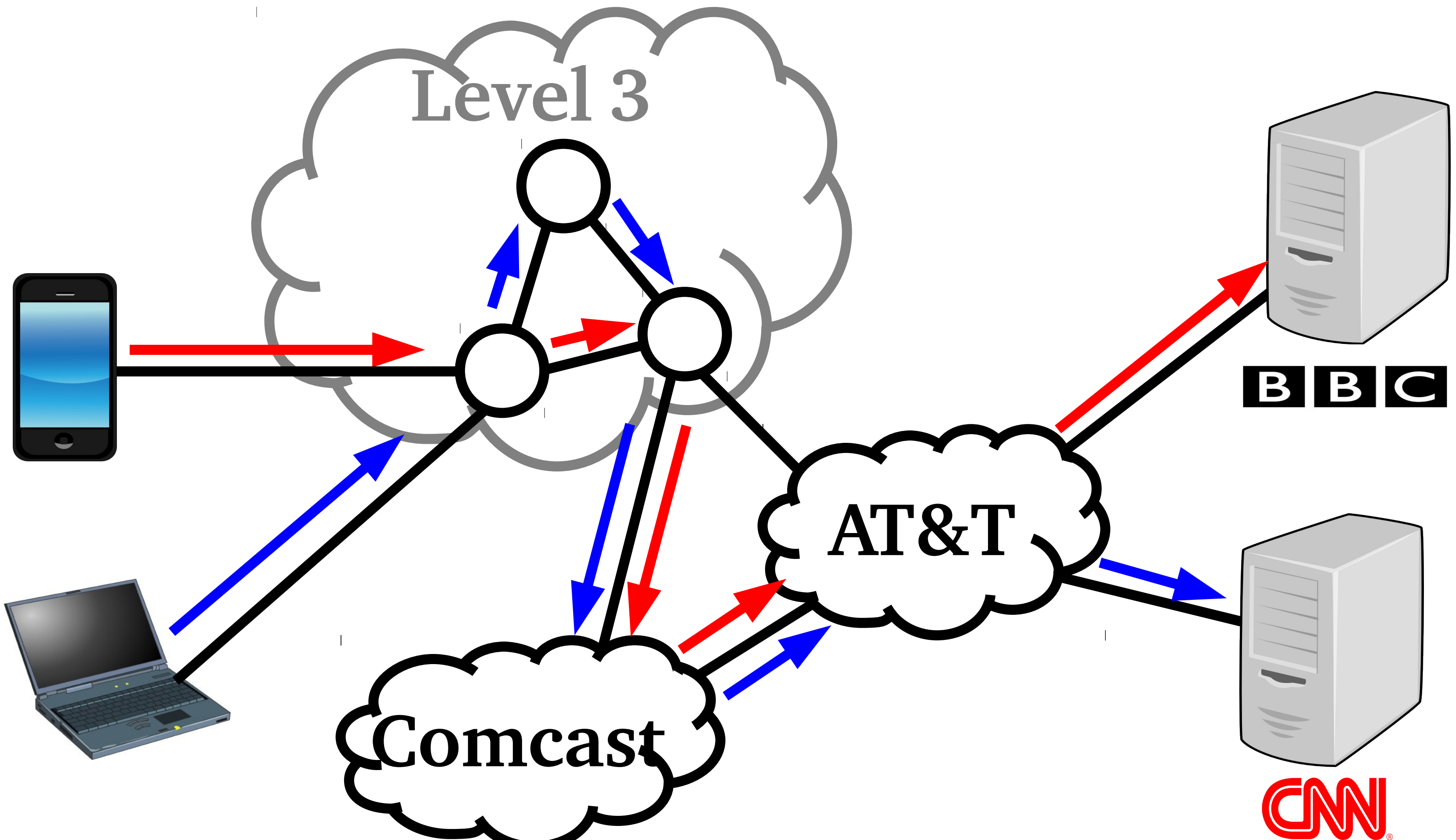
interdomain routing: routing policies

- autonomous entities (Internet Service Providers, ISP)
- distributed (Border Gateway Protocol, BGP)
- economic relationships ($$$)
- consensus problem



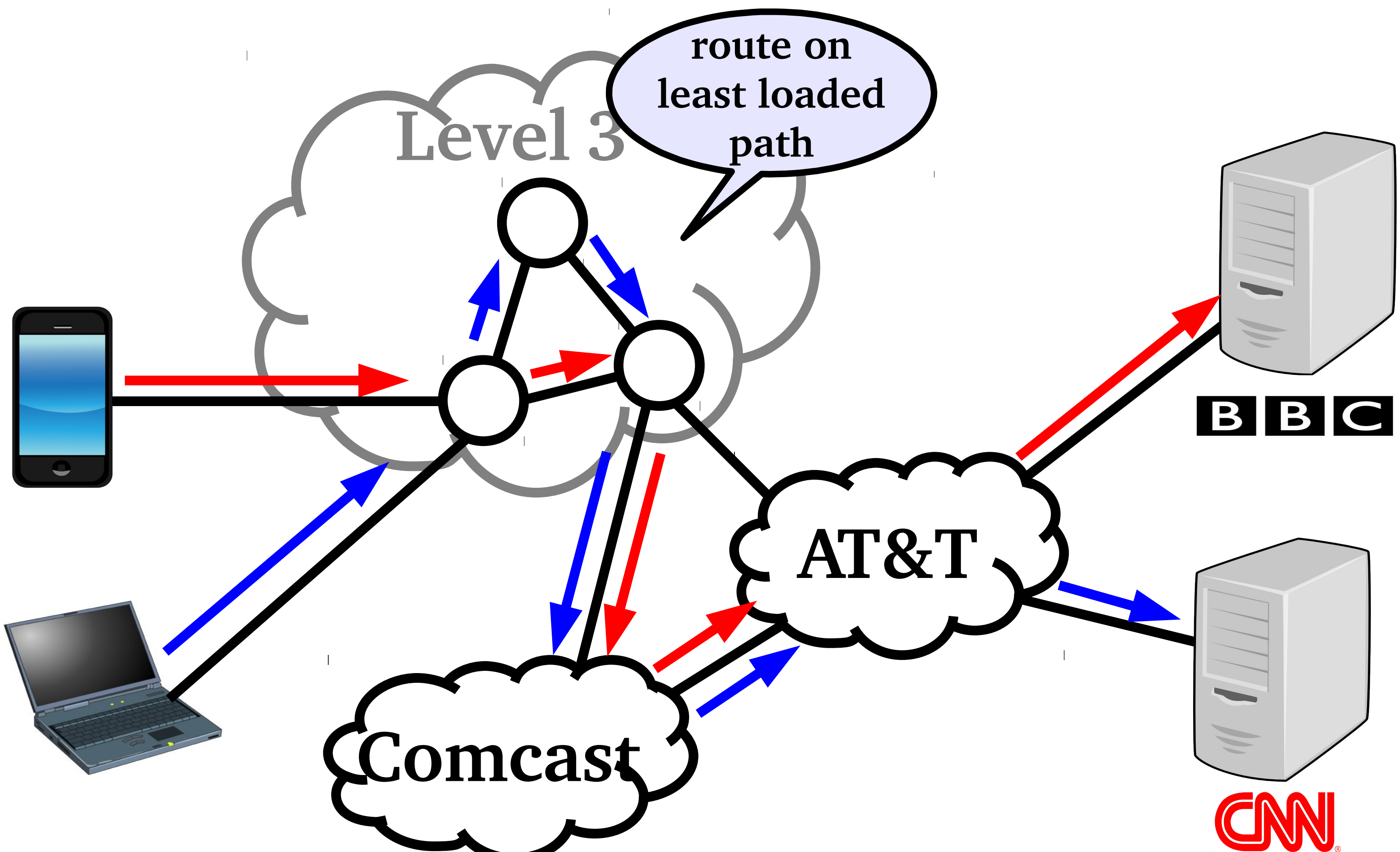
intradomain routing

- full visibility and control



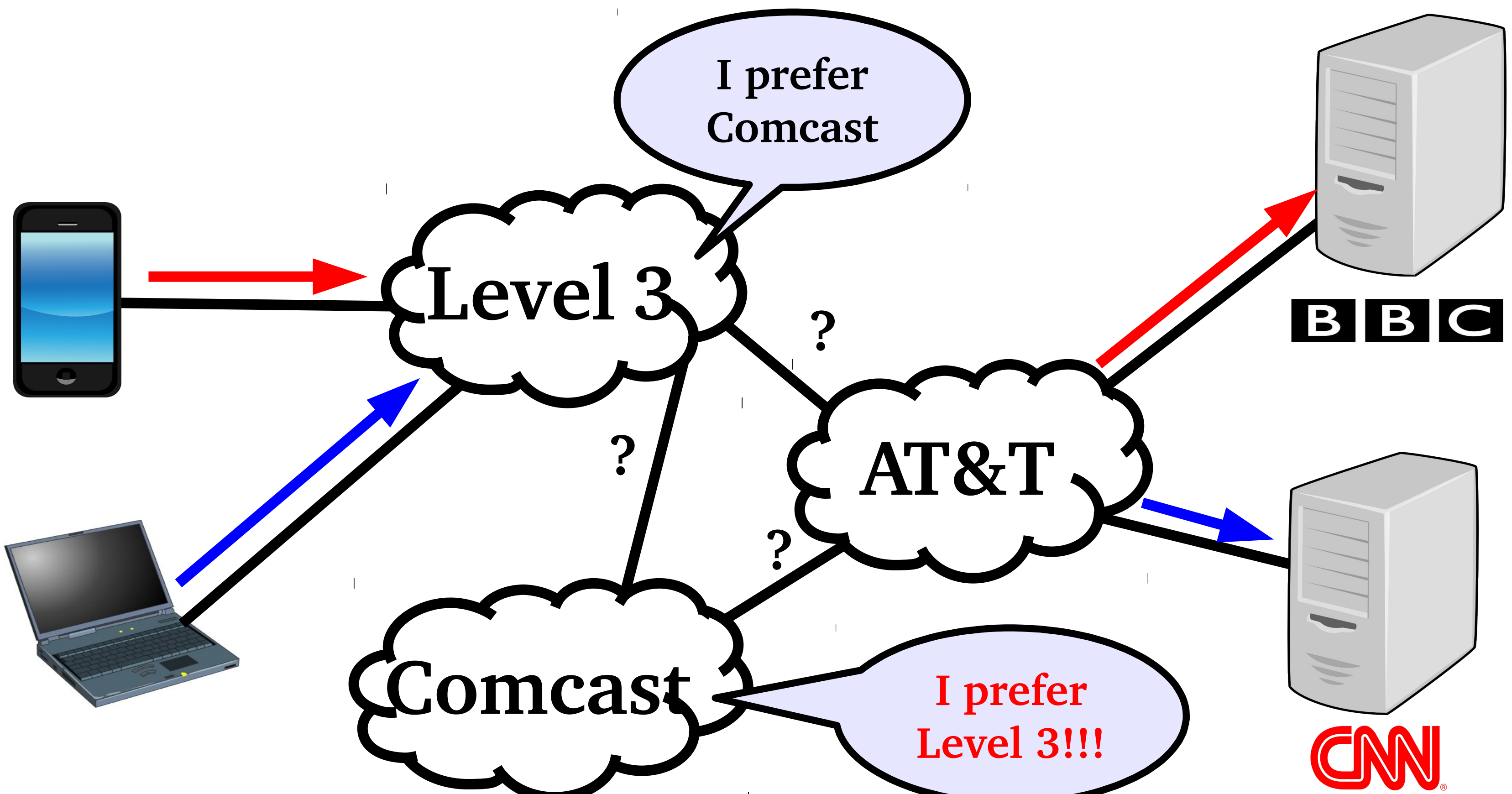
intradomain routing: routing policies

- full visibility and control



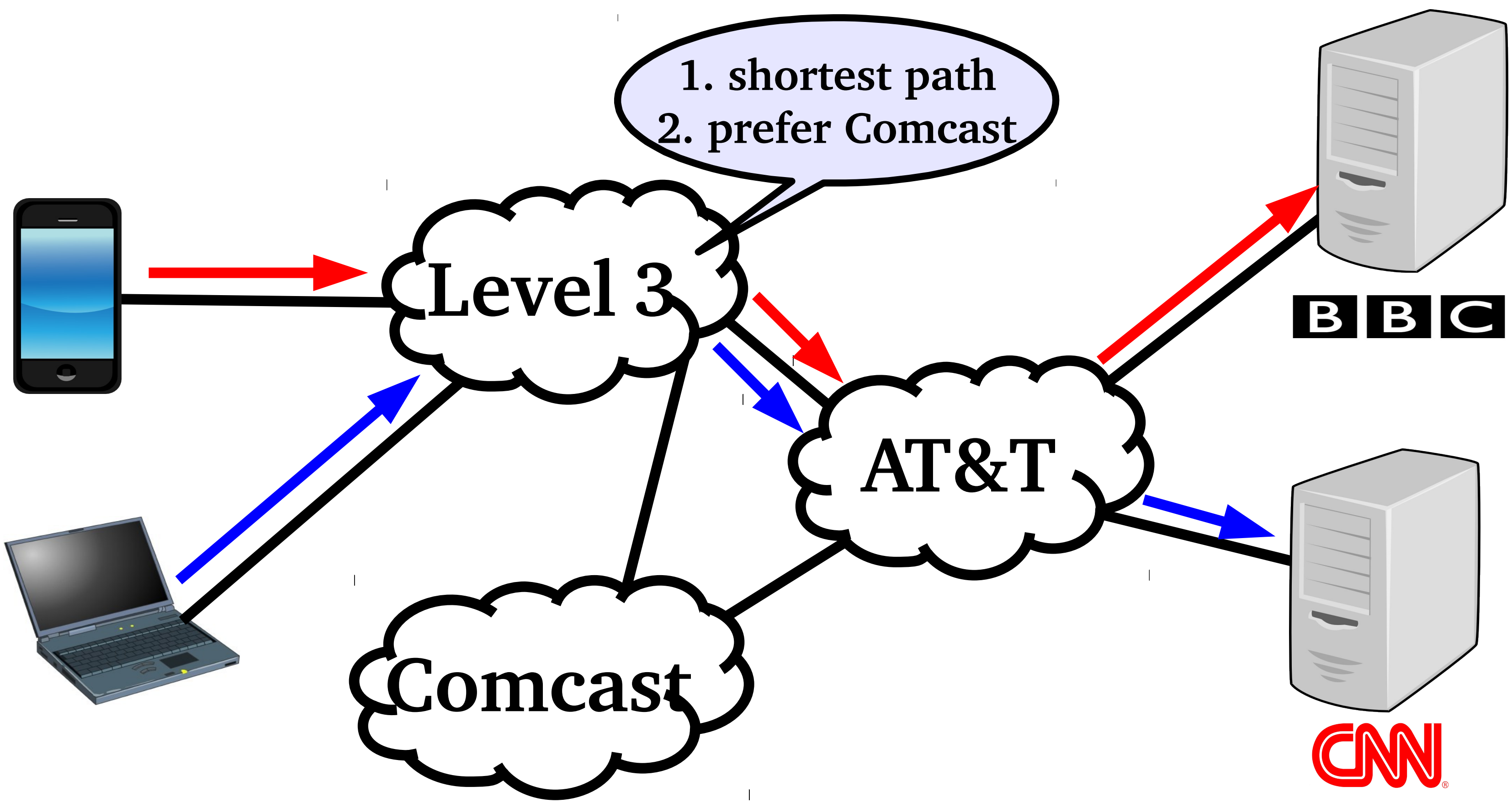
routing challenges (1/3)

stability: are routers guaranteed to agree on a specific routing?



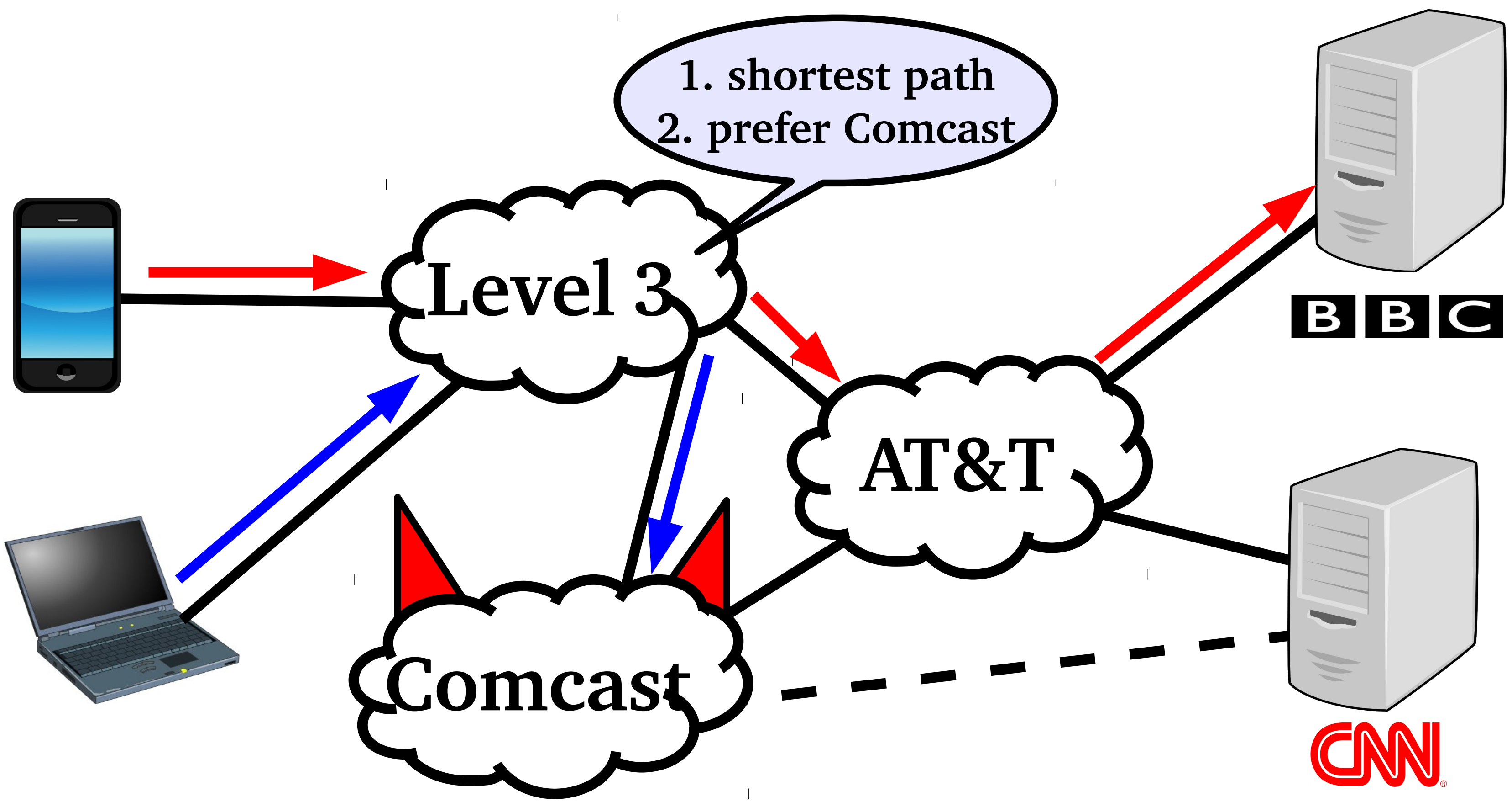
routing challenges (2/3)

security: how do local (truthful/bogus) routing changes influence global routing?



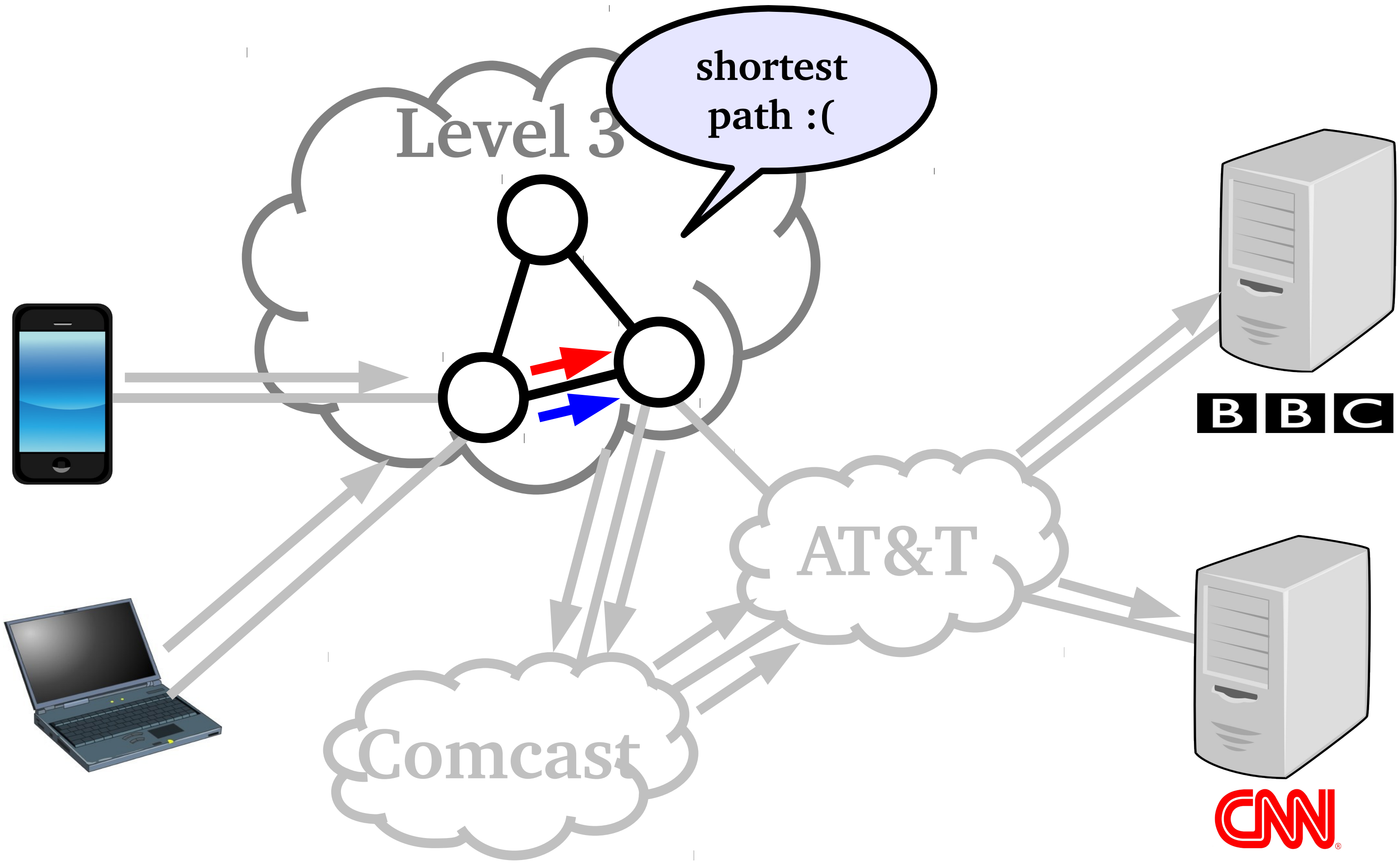
routing challenges (2/3)

security: how do local (truthful/bogus) routing changes influence global routing?



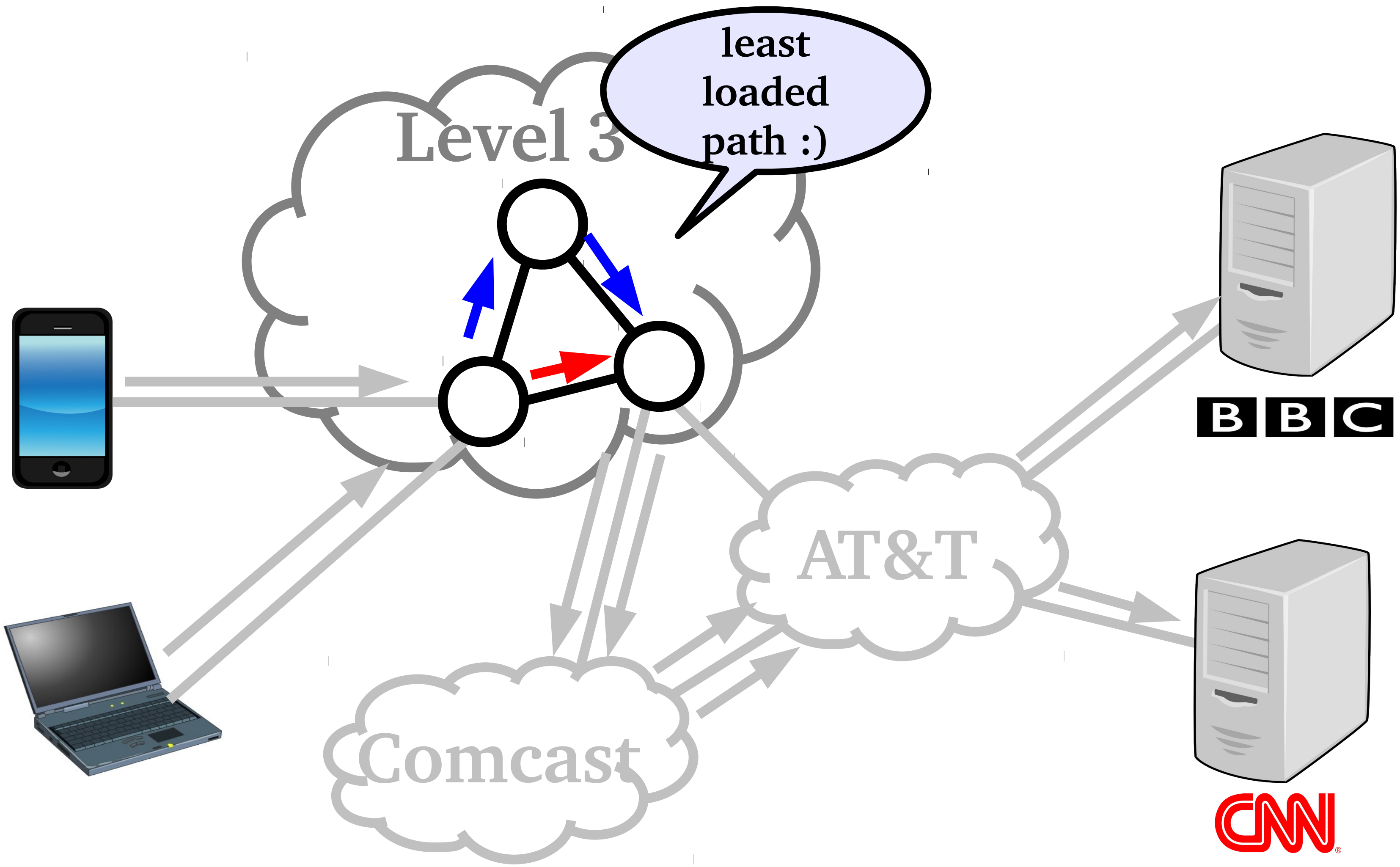
routing challenges (3/3)

load-balancing: how to maximize network resources utilization?



routing challenges (3/3)

load-balancing: how to maximize network resources utilization?



contributions

arbitrary topologies +
arbitrary routing policies =

---

computationally intractable

contributions

arbitrary topologies +
arbitrary routing policies =

---

computationally intractable

we show how to achieve
computational tractability
by restricting
policy expressiveness or
topologies

stability: expressiveness of interdomain routing policies

goal: reach consensus on a stable routing

routing tables are computed in a distributed way:

- receive route announcements from your neighbors

stability: expressiveness of interdomain routing policies

goal: reach consensus on a stable routing

routing tables are computed in a distributed way:

- receive route announcements from your neighbors
- choose your best route ← routing policies

stability: expressiveness of interdomain routing policies

goal: reach consensus on a stable routing

routing tables are computed in a distributed way:

- receive route announcements from your neighbors
- choose your best route ← routing policies
- announce it to (some of) your neighbors ← routing policies
- reiterate

stability: expressiveness of real-world interdomain routing policies

ranking:

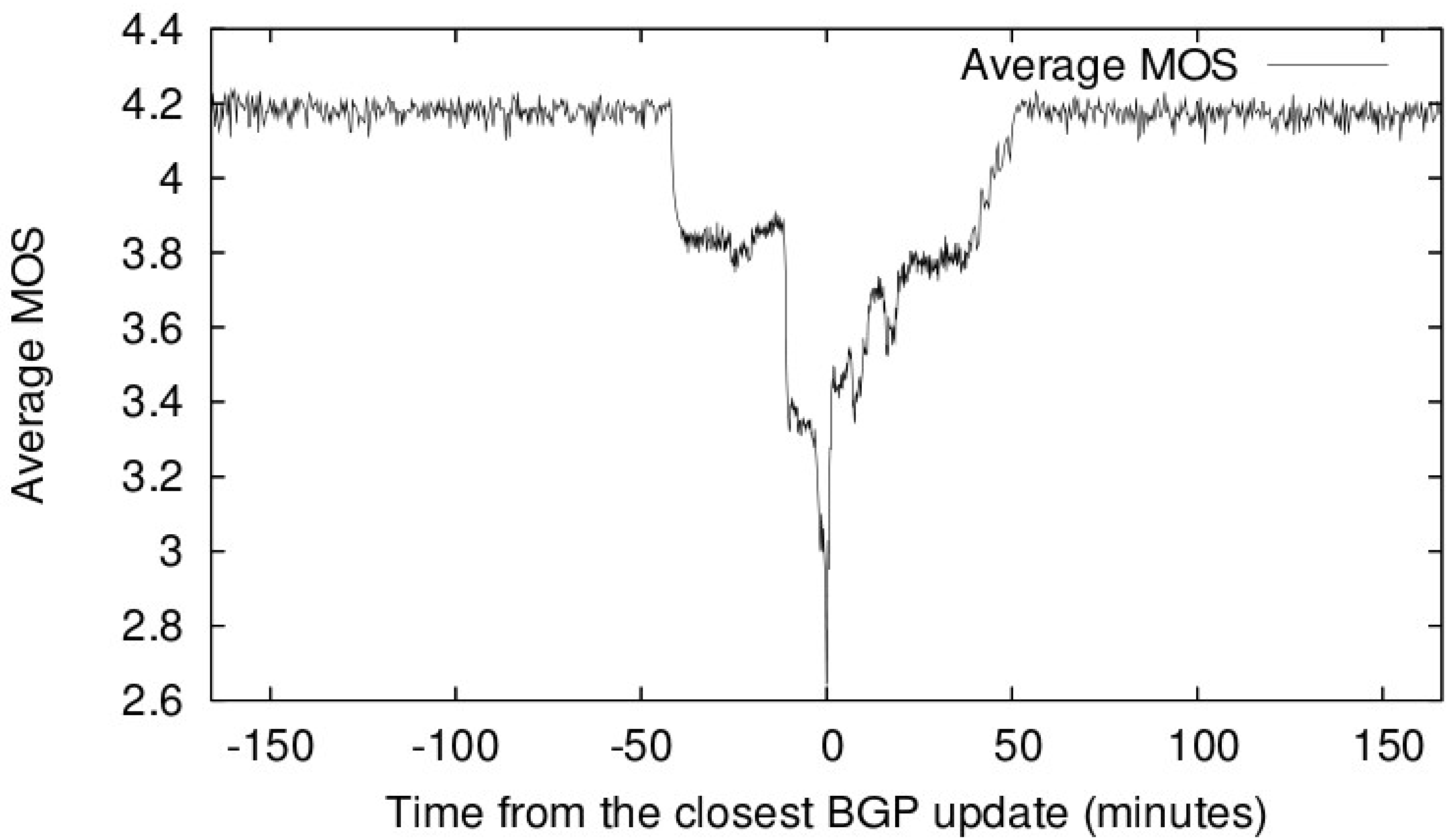
- per-neighbor
- shortest-path

filtering:

- per-neighbor

stability: motivations

- routing is prone to oscillations [Varadhan et al. 2000]
- unpredictable routes propagation [Griffin et al. 2002]



[Kushman, Kandula, Katabi “Can you hear me now?! it must be BGP”, 2007]

stability: problem

can we check if, given a set of routing policies, routers are guaranteed to agree on a specific routing?

known results:

easy for policies with limited expressiveness

shortest-path
Gao-Rexford } → always guaranteed to agree

stability: our contribution
[infocom 2011]

can we check if, given a set of routing policies, routers are guaranteed to agree on a specific routing?

answer: No, computationally intractable

NP-Hard to check it for arbitrary per-neighbor policies

easy only for simple routing policies:

- e.g., filter "all or nothing" per-neighbor

stability: our contribution

[infocom 2011]

can we check if, given a set of routing policies, routers are guaranteed to agree on a specific routing? **No!**

can we check if they agree in $<n$ steps? **No!**

can we check for robustness? **No!**

can we check for well-known sufficient conditions? **No!**

stability: our contribution

[infocom 2011]

can we check if, given a set of routing policies, routers are guaranteed to agree on a specific routing? **No!**

can we check if they agree in $<n$ steps? **No!**

can we check for robustness? **No!**

can we check for well-known sufficient conditions? **No!**

every interesting problem is computationally hard

how hard are them?

why are they so difficult?

stability: a novel mapping between routing policies and logic circuits

[icnp 2013]

basic idea:

- simulate logic gates by ranking and filtering

stability: a novel mapping between routing policies and logic circuits

[icnp 2013]

basic idea:

- simulate logic gates by ranking and filtering
- routing dynamics ↔ circuit logic dynamics

stability: a novel mapping between routing policies and logic circuits
[icnp 2013]

basic idea:

- simulate logic gates by ranking and filtering
- routing dynamics ↔ circuit logic dynamics
- computational complexity lower bound for circuit logic problems apply to interdomain routing problems.

stability: main result
[icnp 2013]

analyzing interdomain routing dynamics is

as hard as

analyzing a computer program

stability: implications

[icnp 2013]

- no SAT solvers (much harder than many optimization problems)
- ≈ can't predict the routing outcome without letting the system run
- oscillation patterns of exponential length

stability: expressiveness restrictions
[icnp 2013]

a mapping exists even if:

- policies are constrained to satisfy two out of three Gao-Rexford conditions

- policies are “internal BGP” compliant

- routing is based on three simple metrics
(e.g., shortest path, largest bandwidth, reliability)

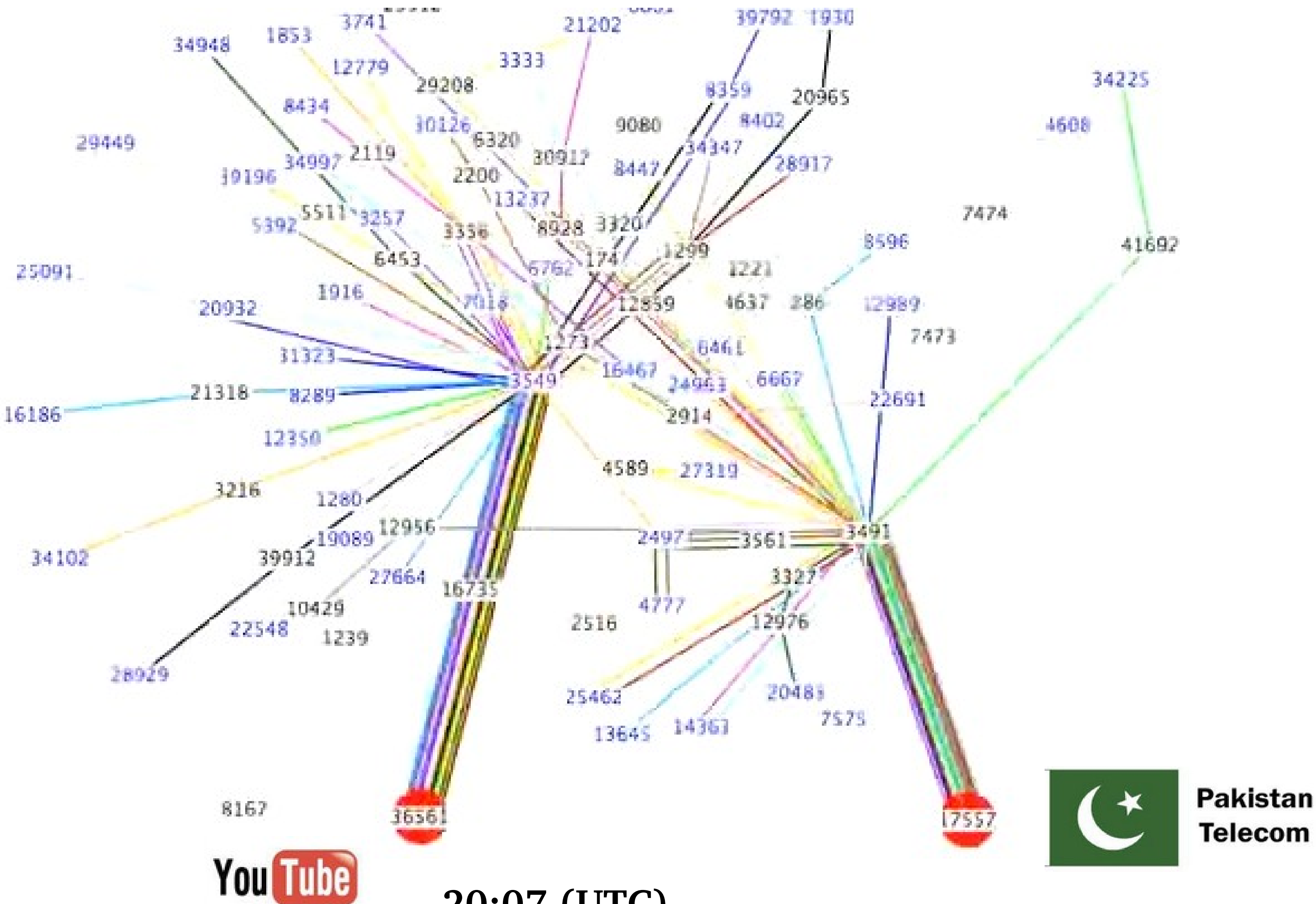
security: how do local changes influence global routing? [icalp 2012]

motivations:

recent attacks on the Internet

security: motivations

[icalp 2012]



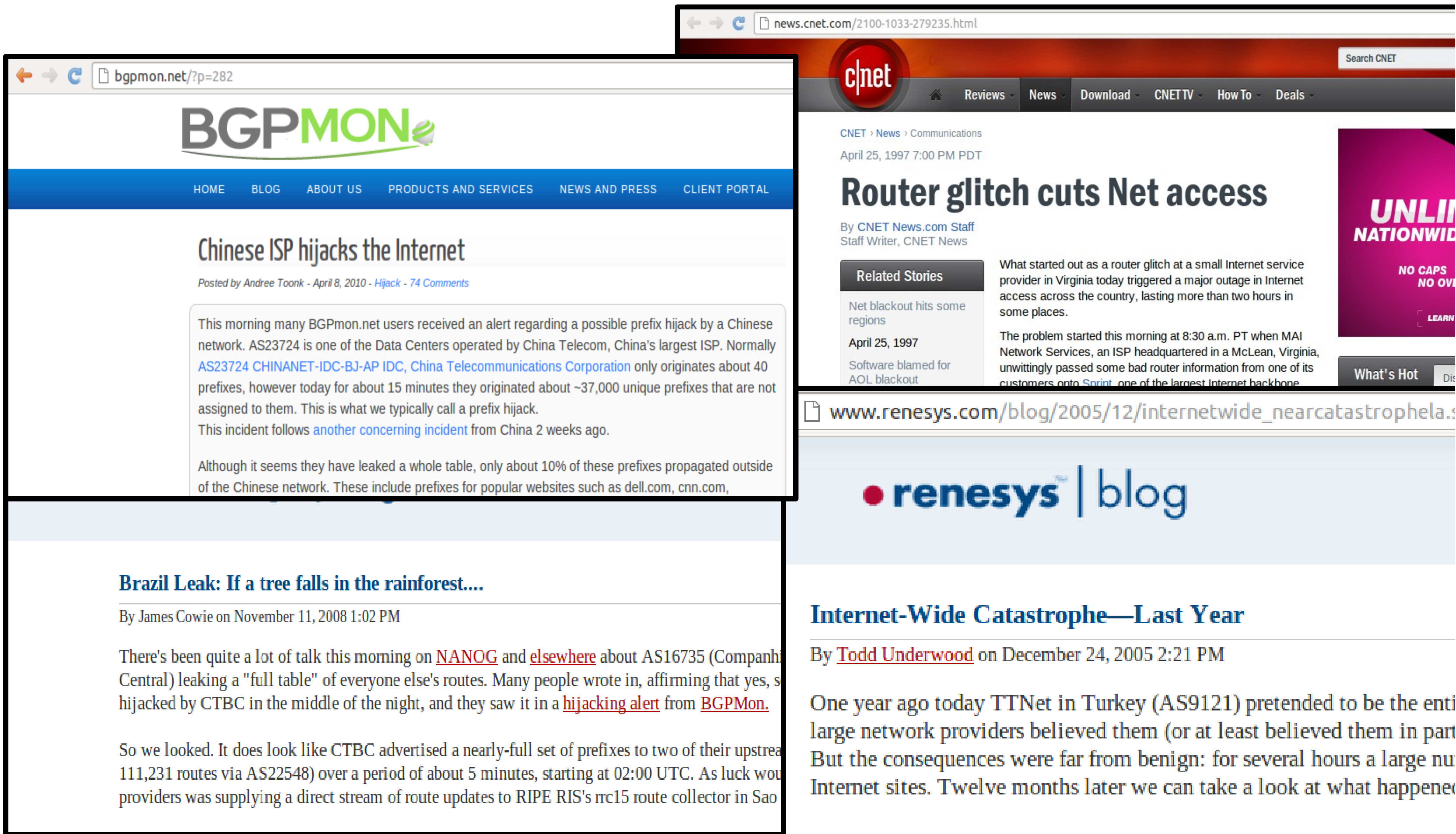
20:07 (UTC)

security: how do local changes influence global routing? [icalp 2012]

motivations:

- recent attacks on the Internet
- possible routers misconfigurations

security: motivations
[icalp 2012]



bgpmon.net/?p=282

BGPMON

HOME BLOG ABOUT US PRODUCTS AND SERVICES NEWS AND PRESS CLIENT PORTAL

Chinese ISP hijacks the Internet

Posted by Andree Toonk - April 8, 2010 - Hijack - 74 Comments

This morning many BGPmon.net users received an alert regarding a possible prefix hijack by a Chinese network. AS23724 is one of the Data Centers operated by China Telecom, China's largest ISP. Normally AS23724 CHINANET-IDC-BJ-AP IDC, China Telecommunications Corporation only originates about 40 prefixes, however today for about 15 minutes they originated about ~37,000 unique prefixes that are not assigned to them. This is what we typically call a prefix hijack.
This incident follows another concerning incident from China 2 weeks ago.

Although it seems they have leaked a whole table, only about 10% of these prefixes propagated outside of the Chinese network. These include prefixes for popular websites such as dell.com, cnn.com,

news.cnet.com/2100-1033-279235.html

cnet Reviews News Download CNET TV How To Deals Search CNET

CNET › News › Communications

April 25, 1997 7:00 PM PDT

Router glitch cuts Net access

By CNET News.com Staff
Staff Writer, CNET News

Related Stories

Net blackout hits some regions

April 25, 1997

Software blamed for AOL blackout

What started out as a router glitch at a small Internet service provider in Virginia today triggered a major outage in Internet access across the country, lasting more than two hours in some places.

The problem started this morning at 8:30 a.m. PT when MAI Network Services, an ISP headquartered in a McLean, Virginia, unwittingly passed some bad router information from one of its customers onto Sprint, one of the largest Internet backbone

What's Hot

Brazil Leak: If a tree falls in the rainforest....

By James Cowie on November 11, 2008 1:02 PM

There's been quite a lot of talk this morning on NANOG and elsewhere about AS16735 (Companhi Central) leaking a "full table" of everyone else's routes. Many people wrote in, affirming that yes, s hijacked by CTBC in the middle of the night, and they saw it in a hijacking alert from BGPMon.

So we looked. It does look like CTBC advertised a nearly-full set of prefixes to two of their upstrea 111,231 routes via AS22548) over a period of about 5 minutes, starting at 02:00 UTC. As luck wou providers was supplying a direct stream of route updates to RIPE RIS's rrc15 route collector in Sao

www.renesys.com/blog/2005/12/internetwide_nearcatastrophela.

renesys | blog

Internet-Wide Catastrophe—Last Year

By Todd Underwood on December 24, 2005 2:21 PM

One year ago today TTNet in Turkey (AS9121) pretended to be the enti large network providers believed them (or at least believed them in part But the consequences were far from benign: for several hours a large nu Internet sites. Twelve months later we can take a look at what happene

security: how do local changes influence global routing? [icalp 2012]

motivations:

- recent attacks on the Internet
- possible routers misconfigurations
- understanding routing vulnerability/predictability

security: how do local changes influence global routing? [icalp 2012]

motivations:

- recent attacks on the Internet
- possible routers misconfigurations
- understanding routing vulnerability/predictability

three questions:

- can I trigger an instability?
- who can hijack my traffic?
- how to intercept traffic?

security: can I trigger an oscillation?

[icalp 2012]

routing policies:

Gao-Rexford

answer: no, every “steady” attack cannot trigger an oscillation

non-steady attacks must be part of the oscillation

security: who can hijack my traffic?

[icalp 2012]

routing policies:

Gao-Rexford

attacks:

origin spoofing → BGP

available-paths → S-BGP

security: who can hijack my traffic?

[icalp 2012]

routing policies:

Gao-Rexford

attacks:

origin spoofing → BGP

available-paths → S-BGP

action space:

deciding to whom neighbor to send a bogus route

security: who can hijack my traffic?

[icalp 2012]

routing policies:

Gao-Rexford

attacks:

origin spoofing → BGP **→ easy to compute**

available-paths → S-BGP **→ hard to compute**

action space:

deciding to whom neighbor to send a bogus route

marks a sharp difference between BGP and S-BGP

security: how to intercept traffic?

[icalp 2012]

routing policies:

Gao-Rexford

answer: announce only one available path

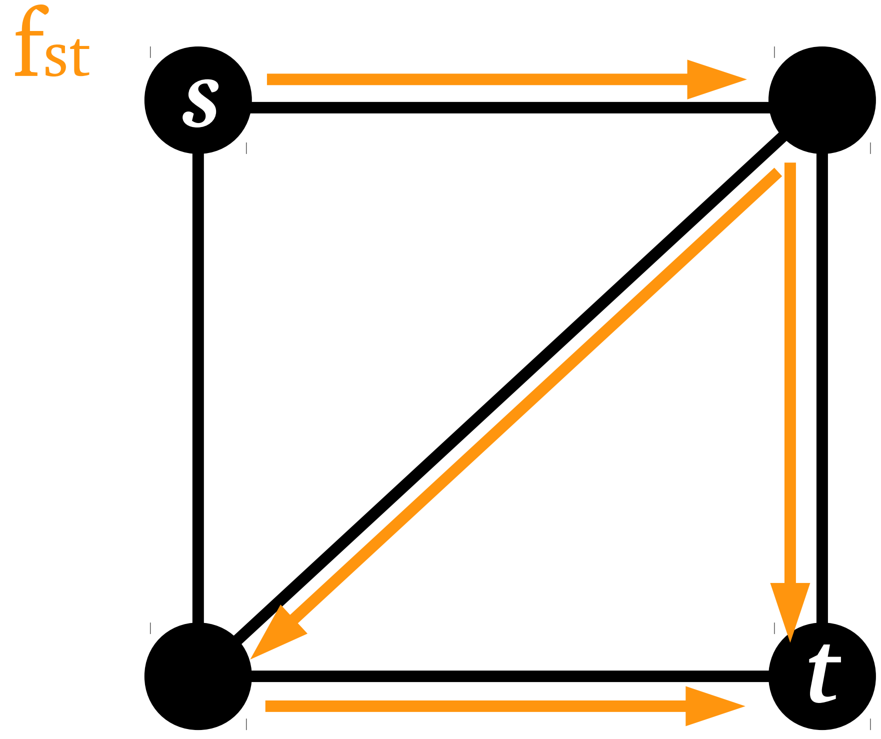
announcing more paths may create “black-holes”

load-balancing: Equal-Split-Max-Flow problem

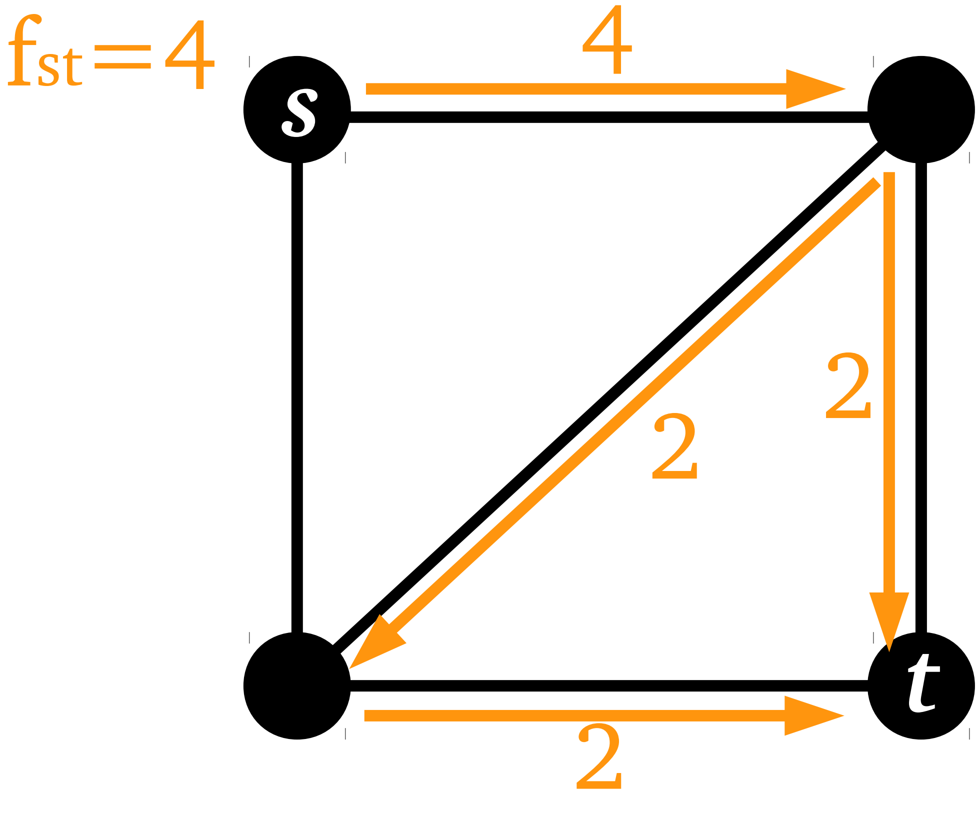
most deployed technique:

- packet header flow-level hash
  - no packet re-ordering
  - if many flows exist → equal-split

equal-split example

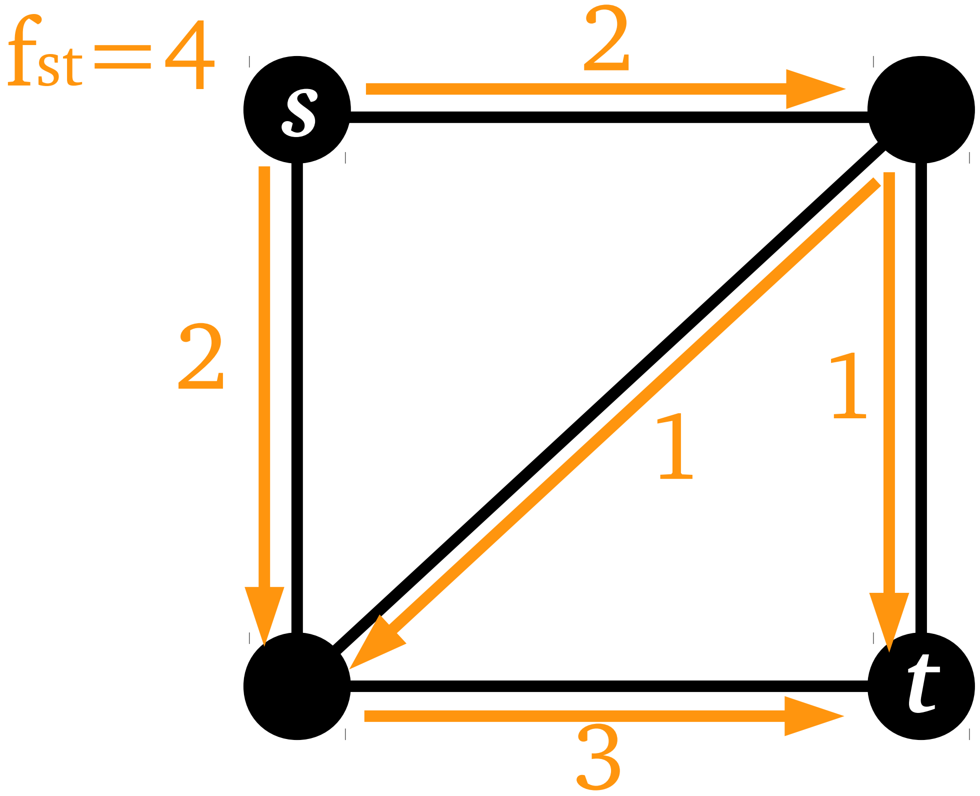


equal-split example



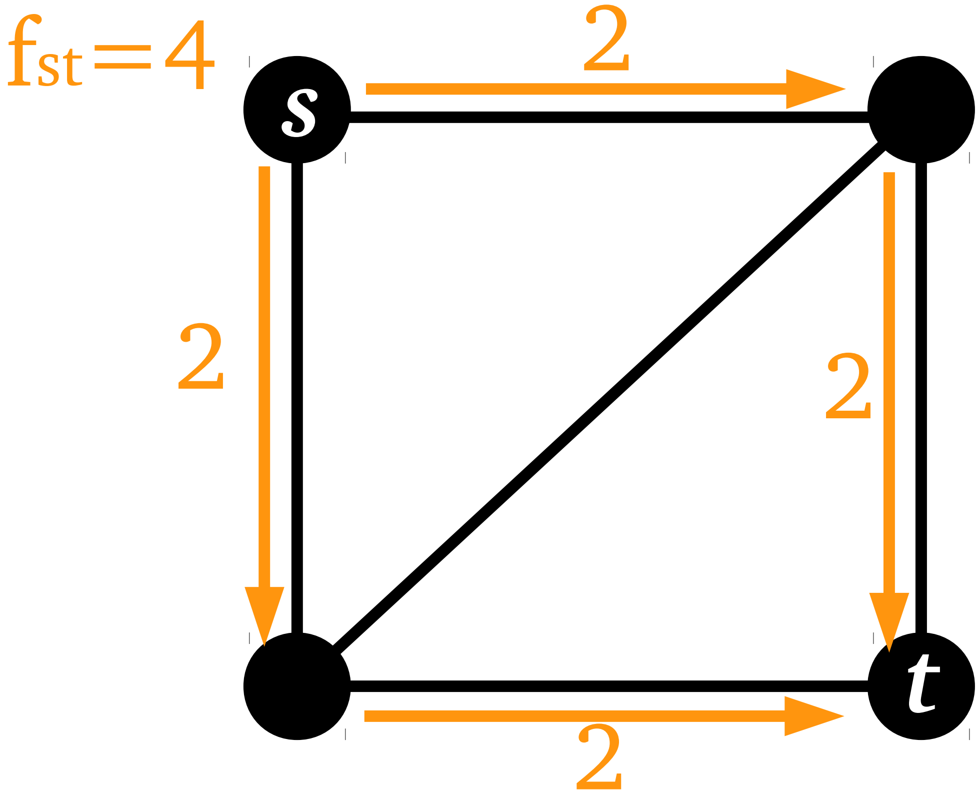
load on most loaded link = 4

equal-split example



load on most loaded link = 3

equal-split example



load on most loaded link = 2

load-balancing: Equal-Split-Max-Flow problem

most deployed technique:

- packet header flow-level hash
  - no TCP re-ordering
  - if many flows exist → equal-split

optimization functions:

- maximize throughput across the network
- minimize most congested link
- minimize sum of link costs

wanted: algorithm with ***provable*** guarantees

load-balancing: Equal-Split-Max-Flow is inapproximable for arbitrary topologies **[infocom 2014]**

known result **[Fortz et al 2002]**:

NP-hard to approximate within a factor of *2/3*

real-network utilization is typically 20%.

our contribution:

NP-hard to approximate within **any** constant factor

- new amplification gap technique

load-balancing: key tool
amplification operator X

operator X: instance I → instance I_{new}

such that

$$OPT(I_{new}) = OPT(I)^2$$

load-balancing: amplifying the gap

$OPT(I) = 1$ or $OPT(I) = \frac{2}{3}$

it is NP-hard to distinguish between ***1*** *and* ***~0.6***

load-balancing: amplifying the gap

$OPT(I) = 1$ or $OPT(I) = \frac{2}{3}$

*it is NP-hard to distinguish between **1** and **~0.6***

$OPT(X(I)) = 1$ or $OPT(X(I)) = \frac{4}{9}$

*it is NP-hard to distinguish between **1** and **~0.4***

load-balancing: amplifying the gap

$OPT(I) = 1$ or $OPT(I) = \frac{2}{3}$

it is NP-hard to distinguish between ***1*** *and* ***~0.6***

$OPT(X(I)) = 1$ or $OPT(X(I)) = \frac{4}{9}$

it is NP-hard to distinguish between ***1*** *and* ***~0.4***

$OPT(X^2(I)) = 1$ or $OPT(X^2(I)) = \frac{16}{81}$

it is NP-hard to distinguish between ***1*** *and* ***~0.2***

…

load-balancing: Equal-Split-Max-Flow in data-center (DC) network topologies

[infocom 2014]

d-hypercubes (bCube-like):

- NP-hard to approximate within a factor of $1-1/d$

Clos networks (VL2-like):

- trivial to compute optimal (oblivious) routing
- no need for expressive routing policies
- however …

load-balancing: routing elephants in datacenter networks

… a few large flows exist in datacenter traffic

load-balancing: routing elephants in datacenter networks

… a few large flows exist in datacenter traffic

- non-negligible probability of collision
  between two elephant
  flows

load-balancing: routing elephants in datacenter networks

… a few large flows exist in datacenter traffic

- non-negligible probability of collision between two elephant flows

our contributions:

- (1/2)-inapproximability
- (1/5)-approximation routing algorithm

conclusions

interdomain routing:

- routing expressiveness and feasibility of stability testing
- mapping technique: logic circuits
- local changes, routing predictability

intradomain routing:

- network utilization inapproximability
- routing algorithms with provable guarantees in DC
- routing large flows in DC