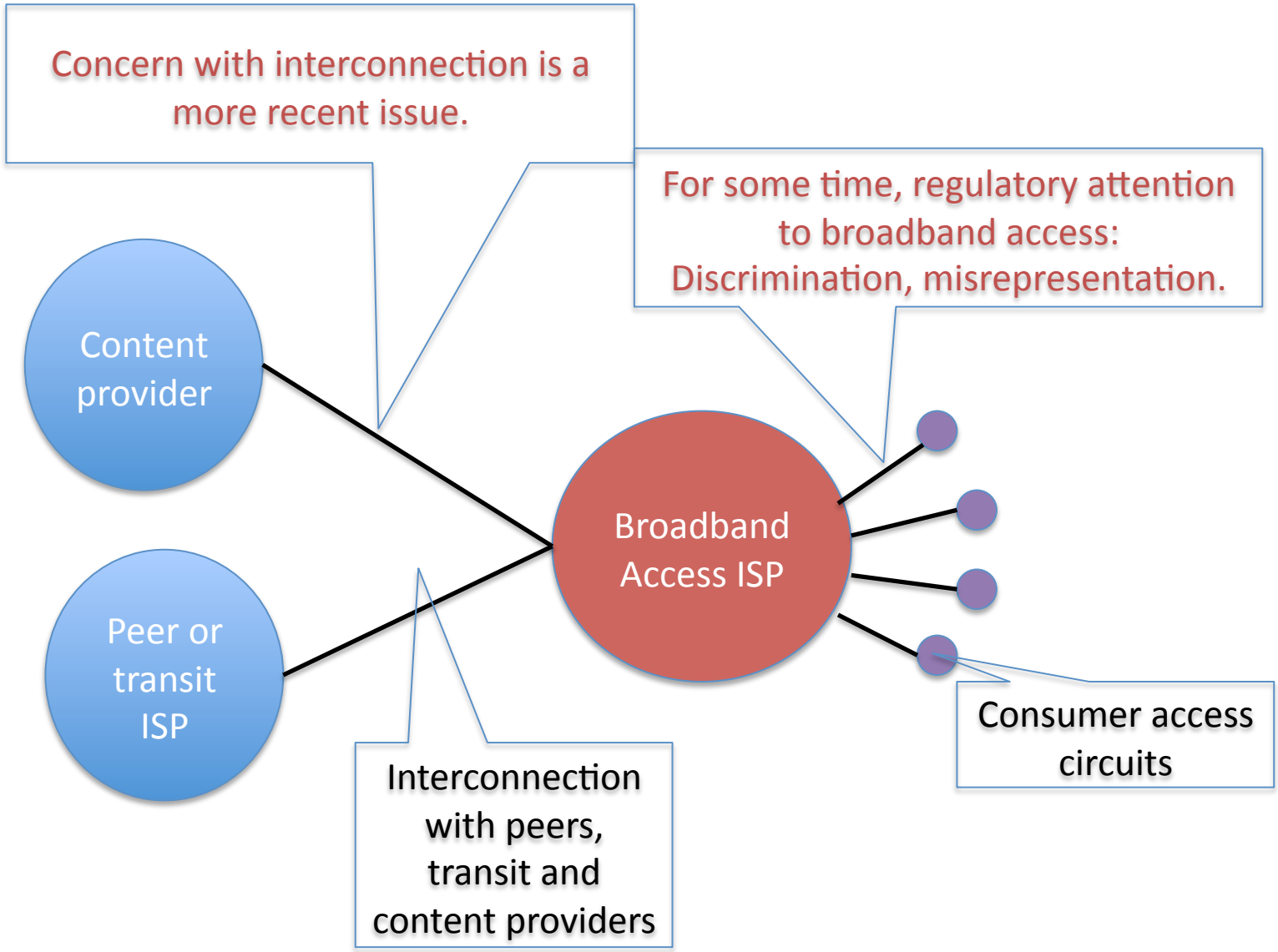
Policy challenges in mapping Internet inter-domain congestion

kc claffy, David D. Clark, Steve Bauer, Amogh Dhamdhere

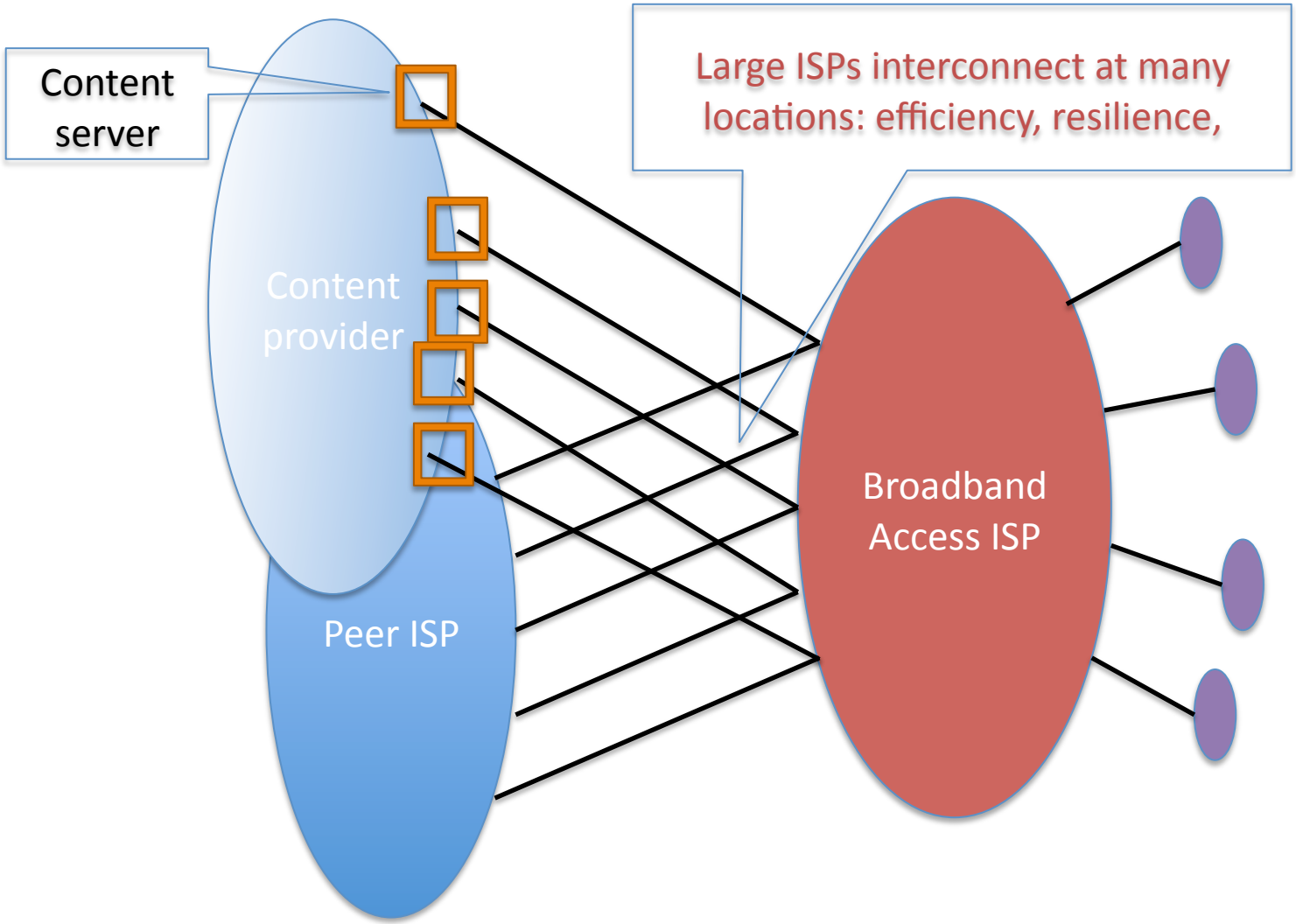
TPRC

September, 2016

Changing focus of attention



A more accurate picture



Some intuition about scale

- The aggregate capacity of interconnection between a large ISP and a large content providers may be several tb/s.
  - Typically distributed among 10 or more locations.
- Typical interconnection links today are 10 gb/s.
  - So there may be hundreds of links at 10 or more points of interconnection.
- Links are grouped into Link Aggregation Groups (LAGs).
  - So a small number of LAGs at a typical interconnection location.
- Trend today: move to 100 gb/s links.

The policy concern

- Points of interconnection might be points for the exercise of market power:
  - Discriminatory pricing or business terms.
    - Or just monopoly rents.
  - Inadequate capacity.
    - Leads to congestion and impairment of the quality of the user experience (QoE).
  - Discriminatory treatment of traffic across the link.
    - Selective dropping or rate limiting.
    - (Not actually the most relevant concern.)

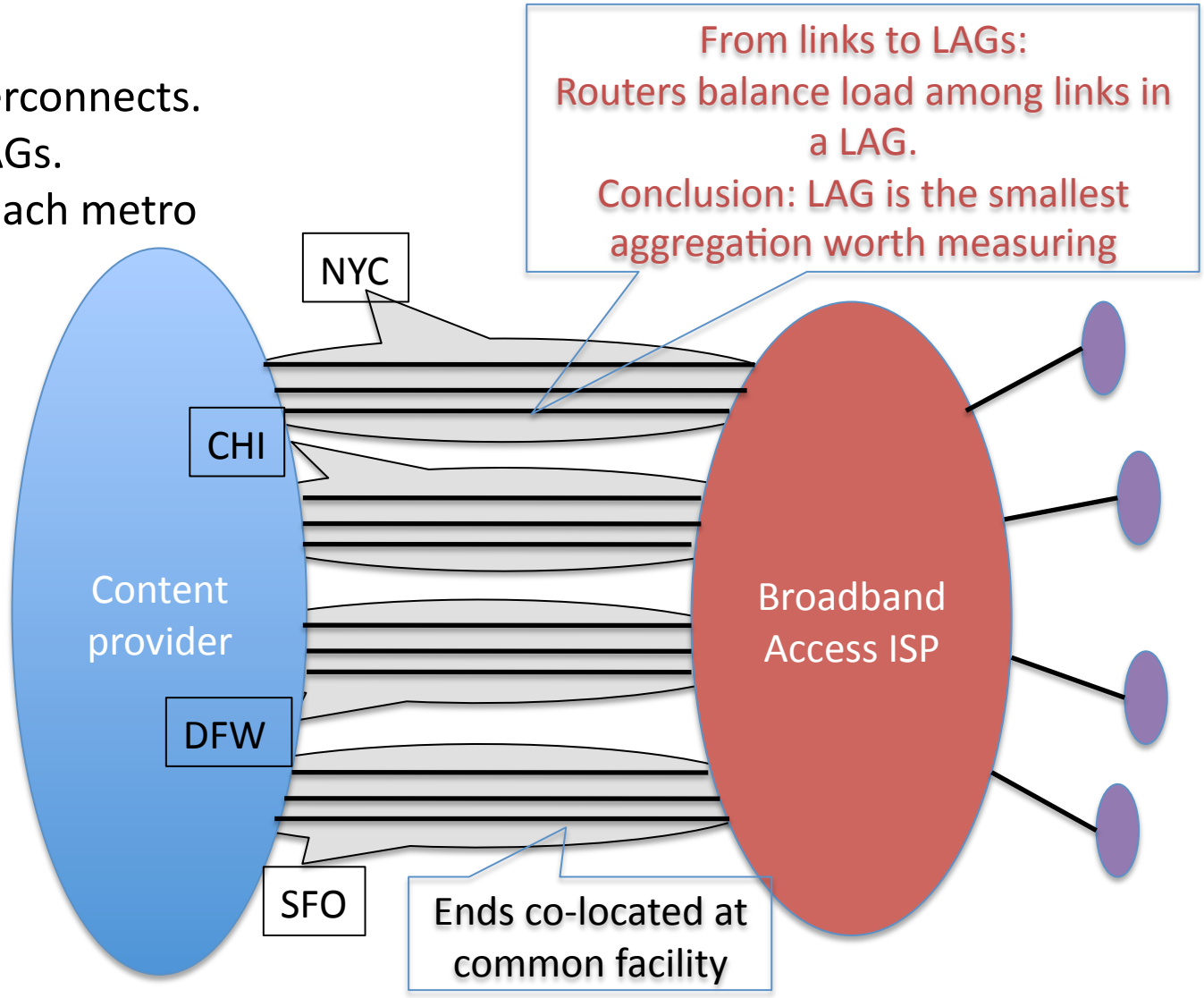
Obvious regulatory response:

- Measure the key characteristics of interconnection links.
  - Or mandate the reporting of those parameters.
- But that response begs two questions:
  - Is measuring individual links actually the right approach?
  - How does one measure the key characteristics of a link?
- This paper explores those two questions.

Metro-area LAGs

In this cartoon:
Four metro interconnects.
The lines are LAGs.
Three LAGs at each metro area.

Is it necessary to report on the individual LAGs in a metro area?



Multiple LAGs in a metro area

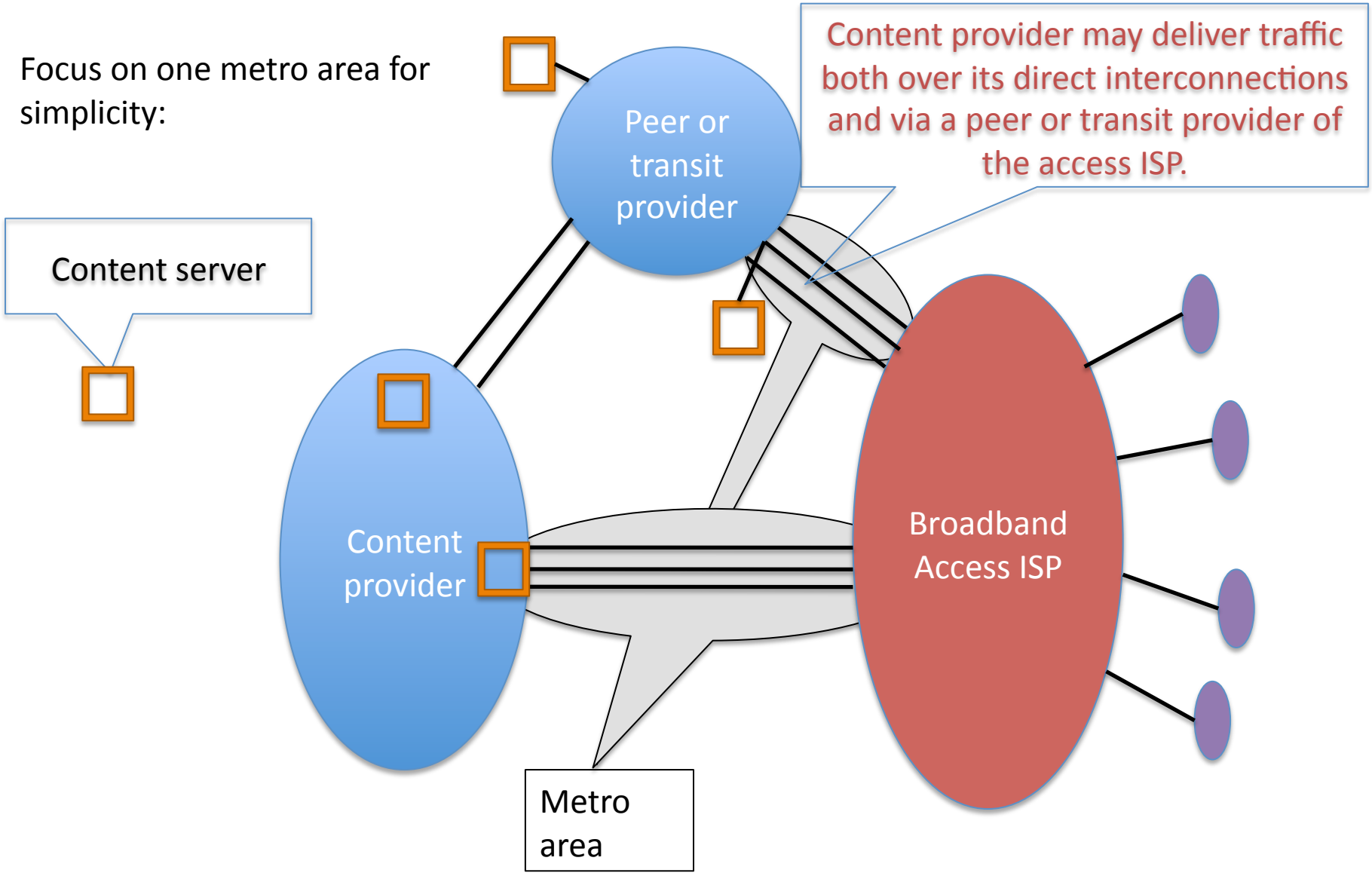
- Questions worth asking:
  - Are they technically substitutable?
    - Give equal performance.
  - Are the business terms similar?
    - Is there a reason to prefer one to another?
- Measurement and reporting:
  - Are the LAGs in a metro area similarly utilized?
    - If so, perhaps only report aggregate utilization.
    - But need measurements of LAGS to confirm that.
- In practice:
  - Imbalanced utilization of LAGs in a metro area happens.

Several levels of aggregation

- Individual links.
  - Too much detail—measure and report on LAGS.
- Metro area LAGs:
  - Aggregate reporting valid if LAGs are substitutable.
    - But must validate substitutability.
  - The “metro” location tells where the interconnection is, not where the customer is.
- Regional area LAGs:
  - Technical definition: region of substitutability.
  - Regulatory definition: region of authority.
- In total between two entities:
  - Expect poor substitutability, so aggregation may mask important variation across the interconnections.

Indirect paths

Focus on one metro area for simplicity:



Indirect paths and aggregation

- When computing the characteristics of some interconnection aggregate (e.g., metro area), should indirect paths (links or LAGs) be included?
  - If actually being used, perhaps yes, but…
    - What entities know if (and which) indirect paths are being used?
    - The content provider knows, the other ISPs may know.
    - Third-party observers cannot usually know.
  - If they are not being used, but are “available”?
    - By what definition of “available”.
    - Again, which actors can know what unused links are practical?

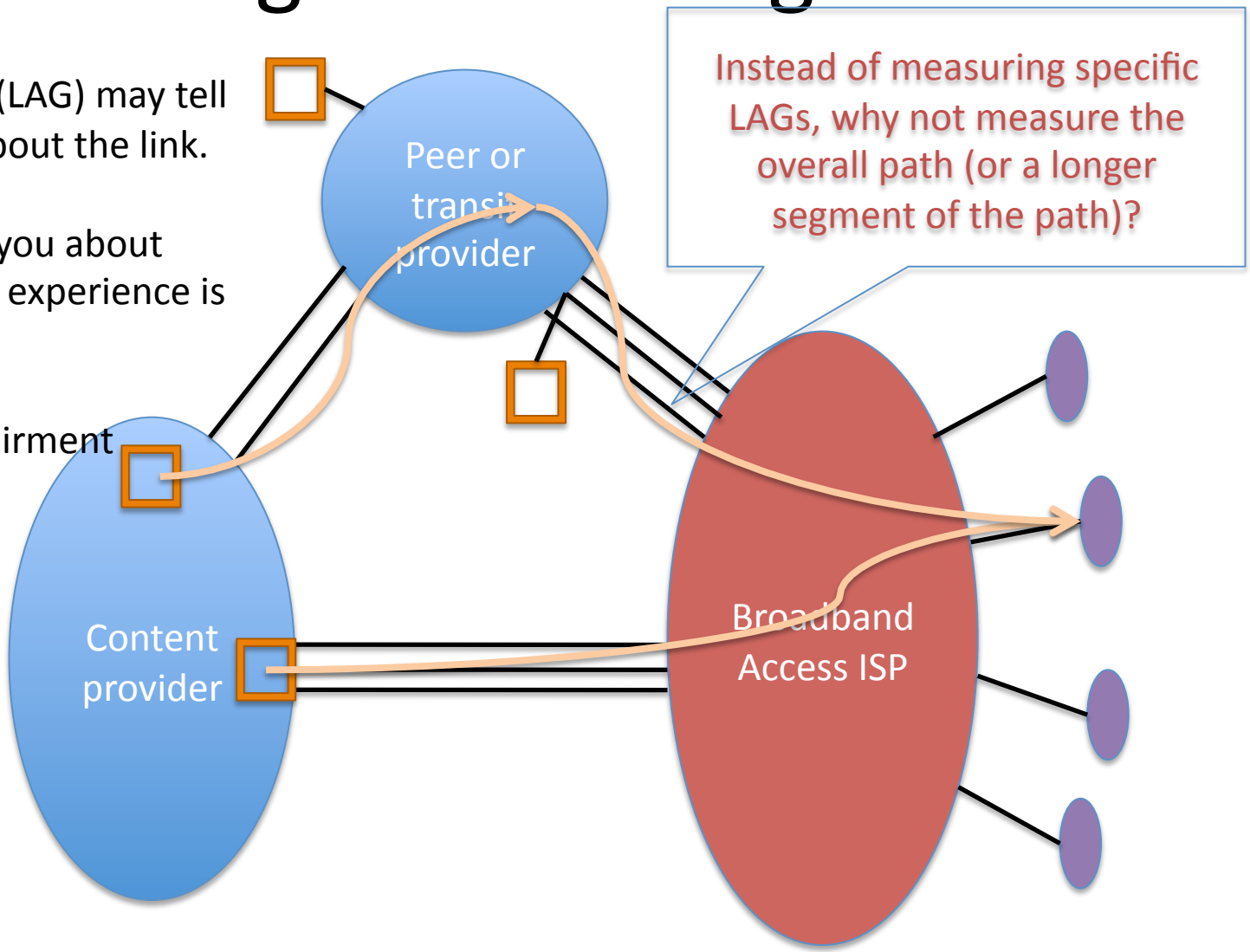
Is measuring LAGs the right idea?

Measuring a link (LAG) may tell you something about the link.

What does it tell you about whether the user experience is impaired?

Perhaps the impairment is elsewhere.

Perhaps there is no impairment.

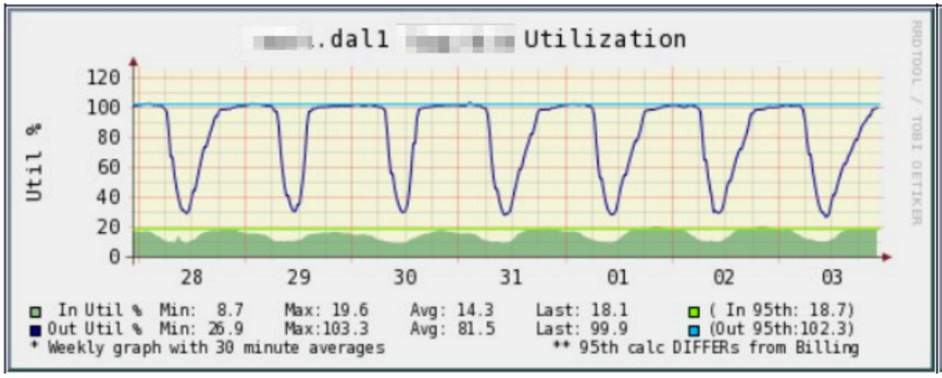


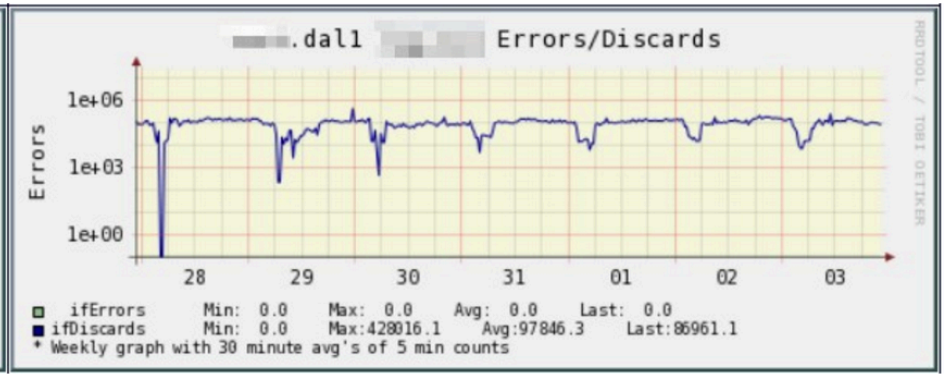
LAGs or paths?

- Measuring the characteristics of a LAG may tell you something about that specific LAG.
  - It will not tell you about whether the user experience is impaired.
- Measuring a path may be a better proxy for a measure of user impairment.
  - But it tells you little about where the impairment happened.
  - Especially, if multiple paths are aggregated.
    - E.g., all direct and indirect paths in a metro area.

The other question

- How to measure the operating condition of a link (LAG)?
- Classic answer: measure utilization.
  - If fully utilized for significant time, assume excess load and congestion.





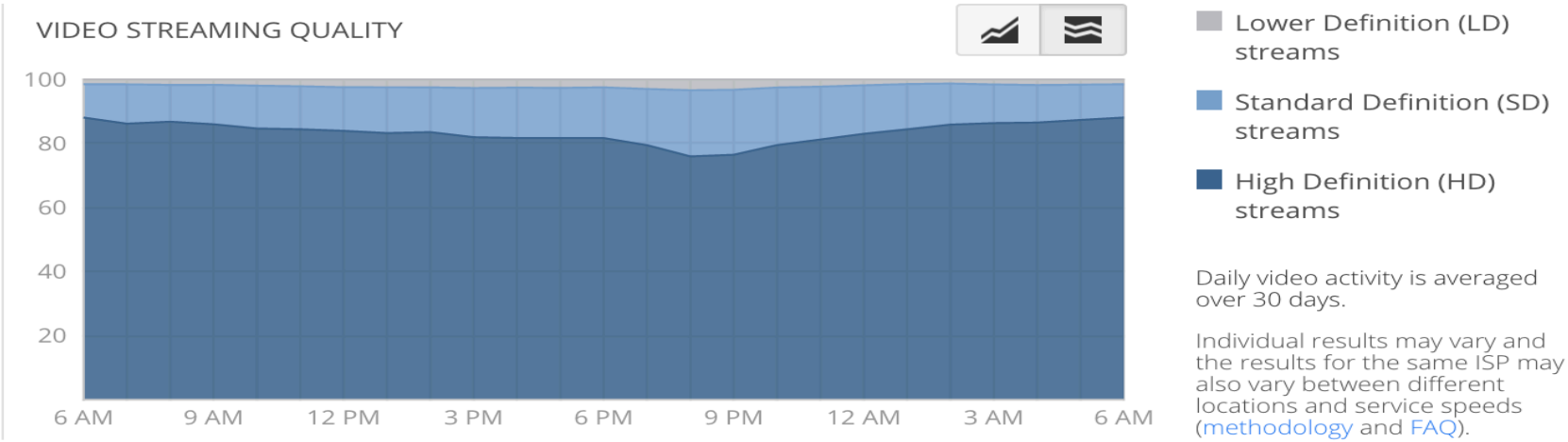
Content providers change the story

- With content providers, the classic answer need not apply:
  - They have control over the source for any specific content, so they can control the path.
  - Might load a LAG to near full but never trigger congestion or any impairment.
  - Need to look for evidence of actual congestion.
    - Packet losses or variation in delay due to queuing (jitter).
- Need multiple measures of a LAG to characterize it.

Six examples of reporting on interconnection.

- Google video quality reports
- Netflix speed index
- Google Measurement Lab
- CAIDA-MIT study
- Princeton Center for Information Technology Policy
- AT&T-DirecTV merger

Google video quality reports



Summarizes video performance at the metro or state level.

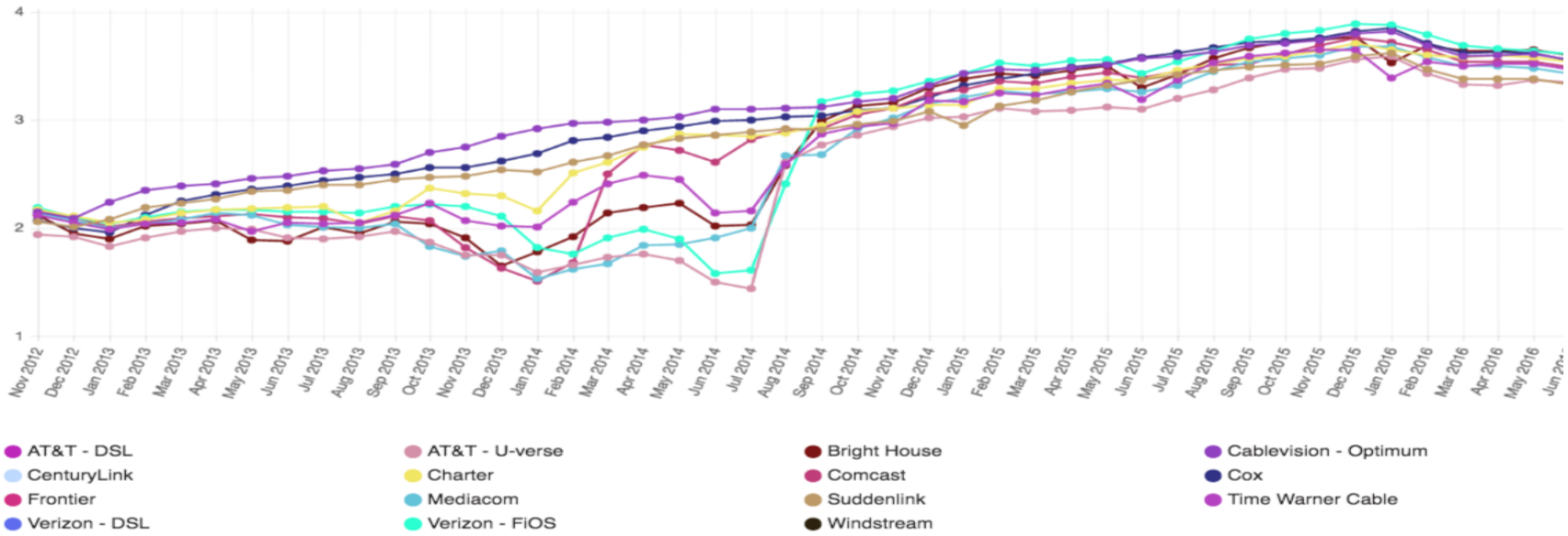
- For a specific ISP in that area.
  - Metro area is defined by where the customer is.

From server to client—end-to-end.

What does it not reveal?

- What paths are included. Do they change over time?
- What caused drop in coding.
- Where any potential impairment is along the path.

Netflix speed index



Aggregate summary of performance across all possible paths by which Netflix serves traffic to a given access provider's customers.

- Monthly data points. No evidence of daily variation.

What does it show?

- Something happened around August 2014.

What does it not reveal?

- Why was there a drop in speed over the last 6 months.
- What is the homogeneity of paths? Includes direct and indirect paths.

Another dimension of reporting

- Over what time granularity?
  - A plot that shows daily variation may suggest peak time congestion.
    - 5 minute intervals? 15?
  - A plot that shows variation over a series of days can show growth in demand, or changes in capacity.
- The choice of granularity again reveals or hides specific indicators.

Google Measurement Lab

- Provides a speed test (NDT) from client to a nearby M-Lab server.
  - For server in a particular peer or transit network, use daily variation in throughput from that server to infer congestion at interconnection point.
    - But why at the point of interconnection?
    - And is the path always the same?
    - Does not reveal this sort of information.

M-Lab Research Team, "ISP Interconnection and its Impact on Consumer Inter-net Performance - A Measurement Lab Consortium Technical Report." http://www.measurementlab.net/publications.

CAIDA-MIT study

- Use probes to near and far side of a LAG to measure round trip delays.
  - Look for variation over a day (peak periods).
  - Infer that increased delay is due to queuing, which implies congestion.
- Cannot measure capacity or utilization.
- Significant challenge is geo-locating LAGs.
  - Are two LAGs in the same metro?

Princeton Center for Information Technology Policy

- Obtained data about utilization of interconnection links from seven ISPs.
  - 5 minute data: utilization, anonymous partner network, metro.
    - What is best way to characterize utilization?
  - For reporting, cannot identify individual metro connections.
    - Specified certain Anonymity Groups (AGs) for reporting.
      - For any ISP (not named), all interconnecting parties.
      - In one metro area, all (if at least 3) ISPs to all parties.
    - We believe that these AGs support very limited useful conclusions.
    - Not clear how to relate utilization to impairment.

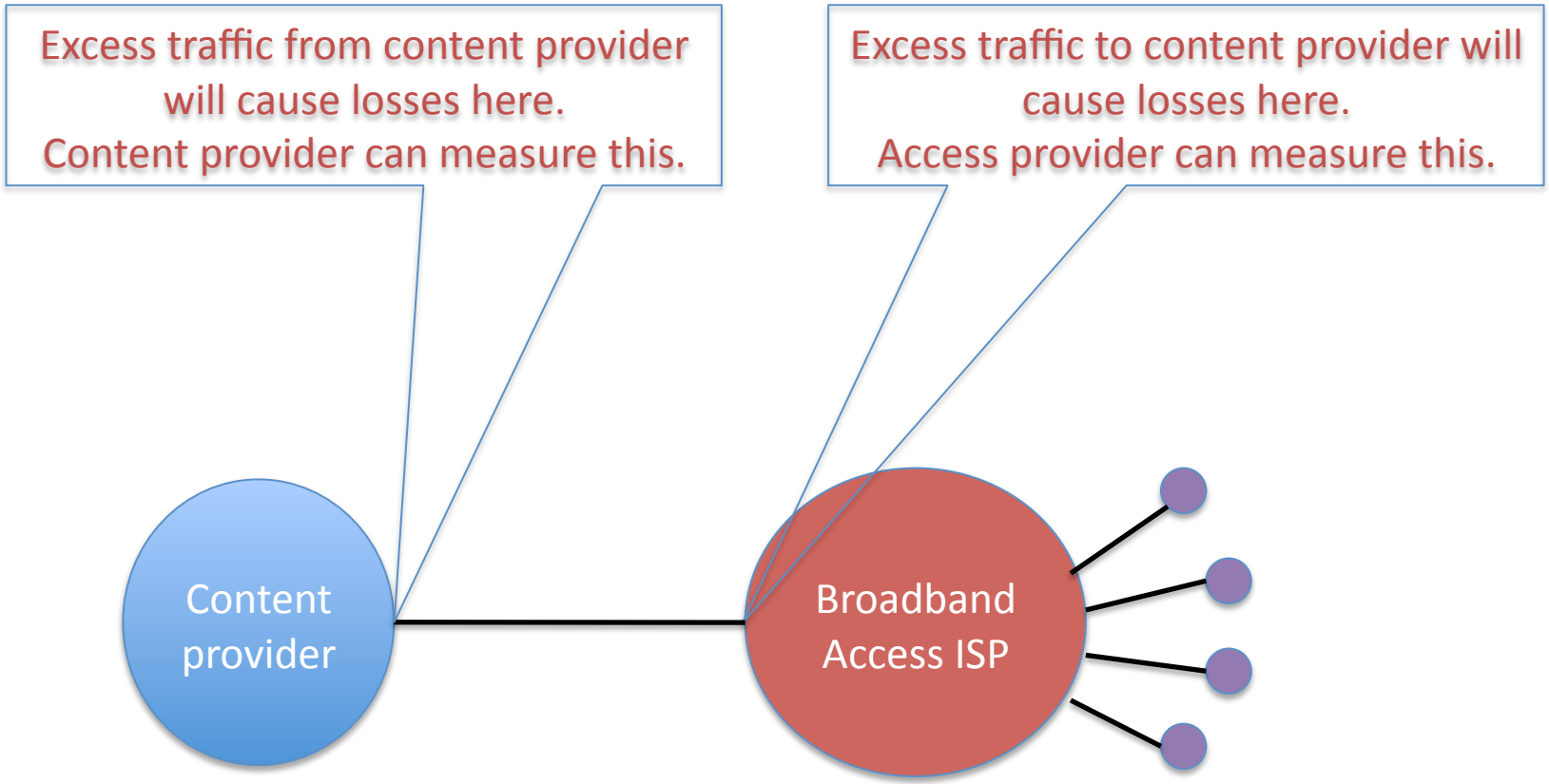
AT&T-DirecTV merger

- We serve as the Independent Measurement Expert to define how AT&T will report on the state of its interconnections with major interconnecting parties.
  - As required by the merger agreement.
  - Order mandates a focus on links.
- Require reporting of utilization, loss and jitter.
  - Report all LAGS, plus metro and total aggregation.
  - But how gather these data points?

Which actors can measure what?

- Either end of a LAG can measure utilization, and presumably knows capacity (and location).
  - Hard (impossible?) for third parties to measure.
- Each end of a connection measures incoming packets that are dropped.
  - But why are they dropped?
  - Each end only sees its incoming traffic.

Where losses happen



This situation caused problems for the ATT reporting requirement. The assumption was that the *incoming* traffic was of more interest, but AT&T could not directly gather data on drops on the "other side" of the interconnection.

Active measurement

- We required AT&T to report on losses two ways?
  - Get data from the other party if possible.
    - Proved difficult in practice.
- Send "probe packets" that solicit a response, and measure percent of these packets that are dropped.
  - By measuring delay, can also assess jitter.
  - But this method is very noisy and unreliable.
    - Massive unexplained losses.
    - Huge variability in delay measures.
    - Can be distorted by differential treatment of traffic classes.
  - There are better methods, but they require cooperation of both parties.

Some conclusions

- Measurement is (often) political.
  - Cooperation from parties only when it benefits.
- Measurement and reporting in ways that selectively reveal or obscure.
- Measurement of individual LAGs does not tell a complete story.
- Reporting of metro or regional aggregates may be appropriate.
  - IF the LAGs are substitutable. But how know that?
- Path measurements can provide a complementary view.
  - But hard to draw robust conclusions from end-to-end measurements.

Further conclusions

- Each stakeholder brings a unique contribution to overall picture.
  - Third parties (e.g., academics) do not have methods to measure some key parameters.
  - ISPs can only see part of the story.
  - Some important measurements require cooperation of both the interconnected parties.
- FCC has data under protective order. How can the research community be a partner in understanding this data?
  - For access, much data is public (MBA). Research community is being excluded from serious analysis of interconnection.