Deconstructing Internet Paths

An Approach for AS-Level Detour Route Disovery

Sing Wang Ho* **Thom Haddow***[1] Jonathan Ledlie†, Moez Draief* Peter Pietzuch*

* Imperial College London
† Nokia Research

April 21, 2009

[1] thaddow@doc.ic.ac.uk

Motivation

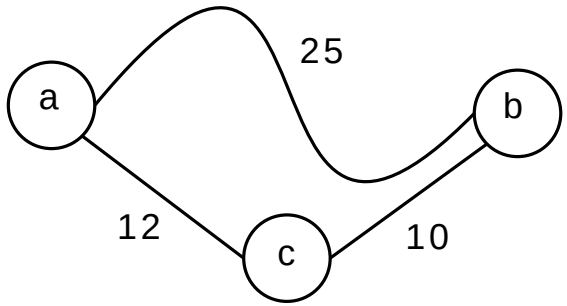
- The Internet provides a *best-effort* end-to-end packet delivery service. QoS guarantees are not generally available.
- Current Internet applications can have wildly differing and dynamic network infrastructure requirements which the existing infrastructure largely does not optimise for.
- It is extremely difficult to adapt the Internet routing infrastructure to be more effective for applications.
- "Best-effort" is a network centric view, but can do better from a user/application perspective by employing **overlay routing**.

Detour Routing

- For any given default Internet path between two end-points there often exists another *potential* path that can offer a better connection quality.
- Since there is no direct means for end-systems to directly control their Internet paths we must instead attempt to exploit these alternate paths through **detour routing**.
- Detour routing is an overlay routing approach which redirects end-to-end traffic via tertiary end-points (**detour nodes**).

TIVs

- For any given cost metric between two end-points a and b, a node c is a detour node for the path $a \rightarrow b$ when the cost metrics between the three nodes form a **triangle inequality violation** (TIV).



- Above, $ab > ac + ab$, thus triangle abc forms a TIV.
- 30–80% of default Internet paths form *significant* TIVs[2]
- Primary causes: Unadvertised connectivity and congestion

[2] [Savage99]

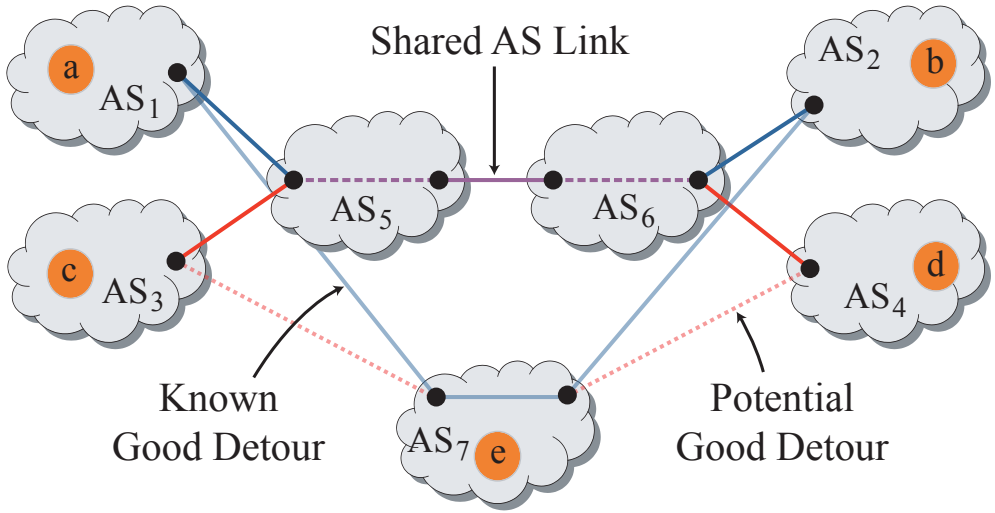
Approach

- TIVs imply the existence of beneficial detour nodes
  - **Finding TIVs enables us to improve on the quality of default Internet paths.**
- Previous work has treated the Internet as a black box[3] but TIVs are often artifacts caused by BGP's operation and routing policies
  - BGP operates at the AS level, thus **the problem is best viewed at this level**
- Most inter-AS connectivity is not publicly exposed by BGP
  - **We must measure this directly**
- Approaches taking too many ($O(n^2)$) measurements do not scale
  - **Any approach needs to take far less measurements**

[3] [RON,Peerwise]

Rationale

- **End-to-end paths traversing similar AS links may benefit from the same sets of detour nodes**



- $path_{a,b}$ has a detour via node e in AS_7
- $path_{c,d}$ is similar to $path_{a,b}$ so may also benefit from detour node e
- **How do we capture and exploit "similarity" between paths?**

Presentation Overview

- Hierarchical clustering of AS paths
- Centralised detour discovery via path clustering
- Decentralised path clustering
- Evaluation of approaches
- Future direction

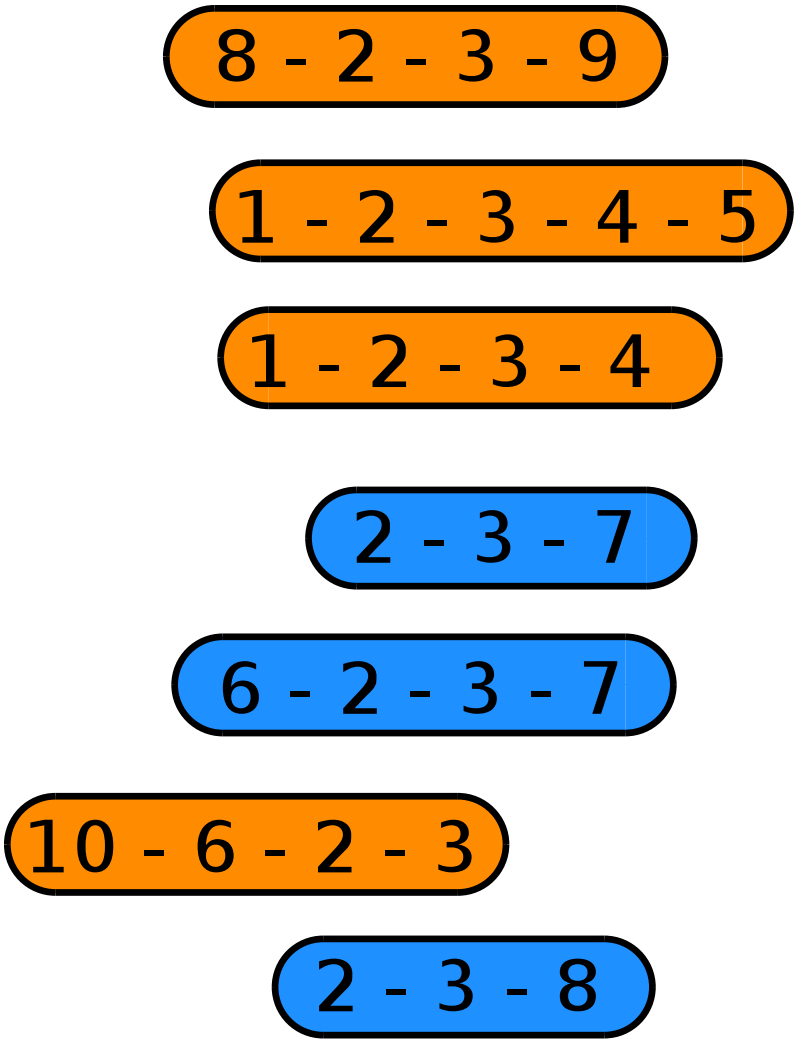
Clustering

- A hierarchical clustering based approach allows us to group similar AS paths into clusters.
- Clustering paths gives us a compact representation which allows us to classify unseen paths based on paths we have already seen.
- Our distance measure for clustering was the **similarity** between AS paths, defined as a function of the number of **shared AS links**.
- From earlier:
  - $path_{a,b} = \{AS_1, \underline{AS_5, AS_6}, AS_2\}$
  - $path_{c,d} = \{AS_3, \underline{AS_5, AS_6}, AS_4\}$

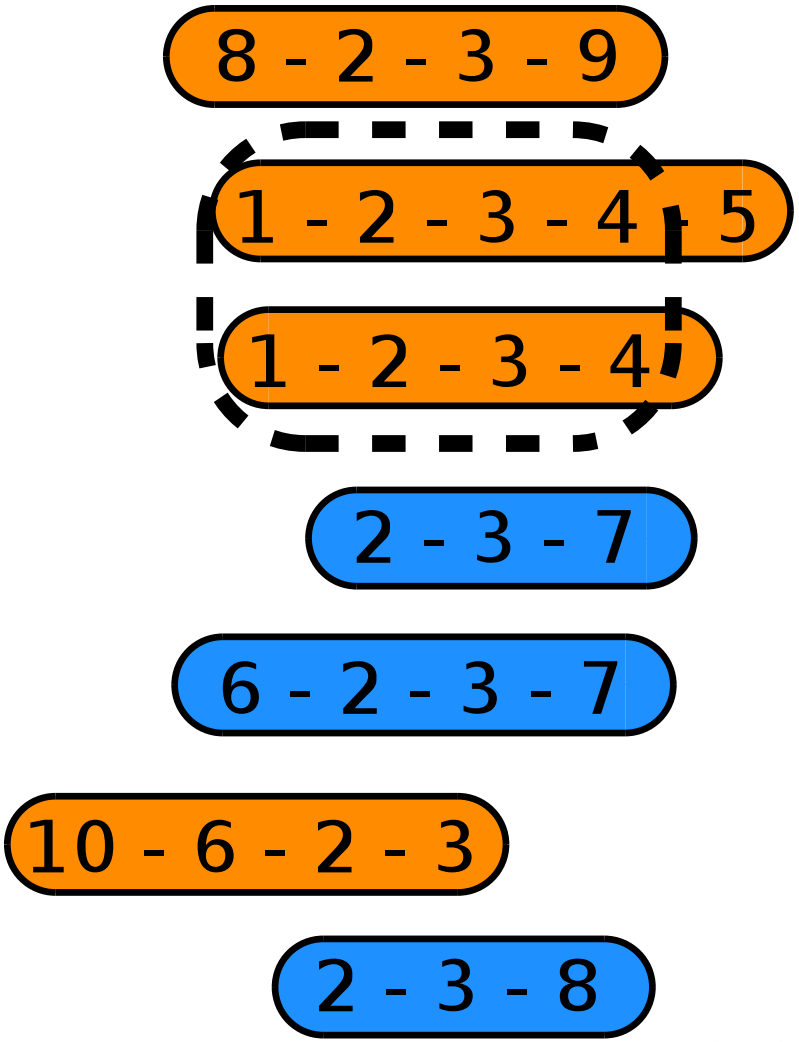
Path Clustering

- We have two path classifications:
  - **TIV paths**: which have a known detour node associated with them
  - **No-TIV paths**: for which no detour node is known
- TIV/No-TIV paths do not mix in clusters
- TIV clusters have a set of associated detour nodes ranked by frequency of occurrence
  - **If a previously unseen path's "closest" cluster is a TIV cluster then the detours in that cluster are likely to be beneficial**
- Clustering is iterative and repeats until no pair of clusters has a similarity value greater than some threshold, as follows...

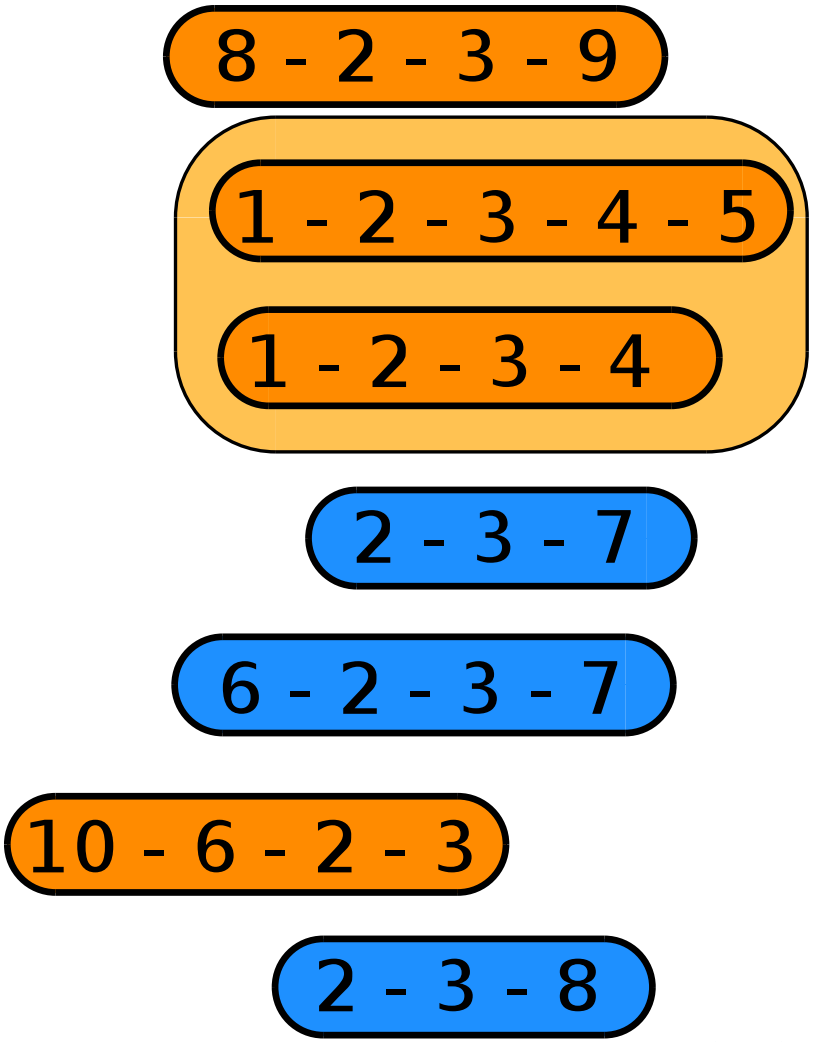
Path Clustering Example



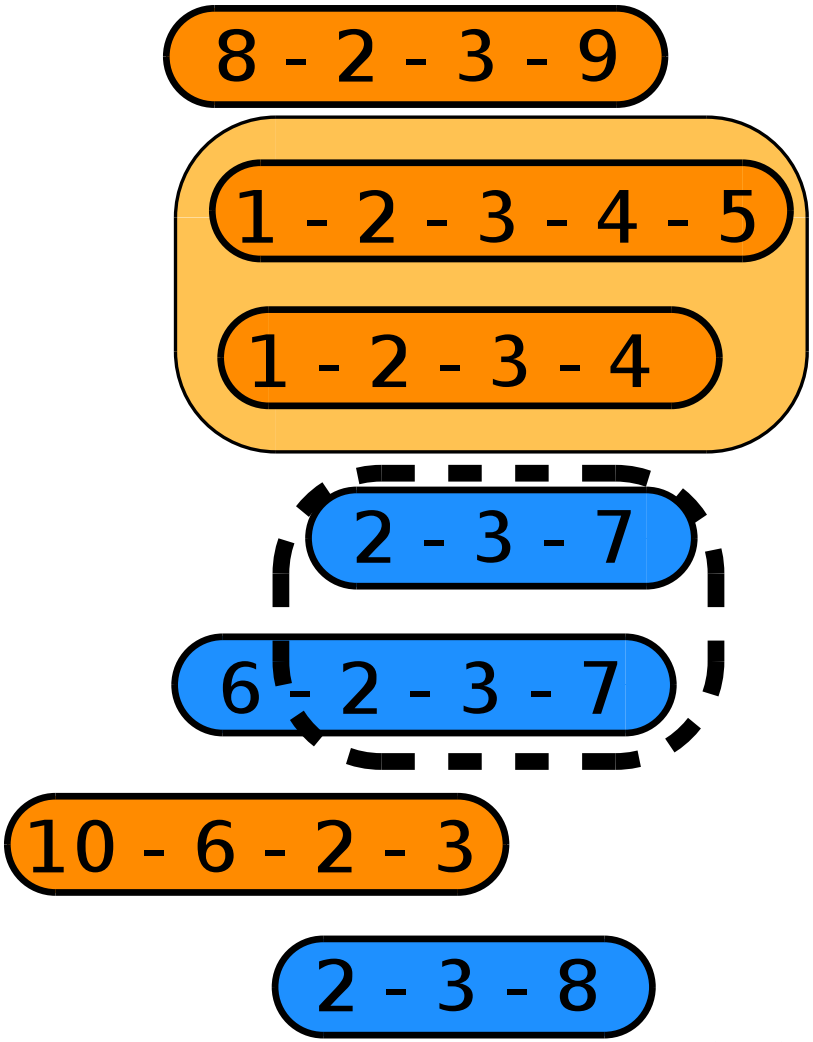
Path Clustering Example



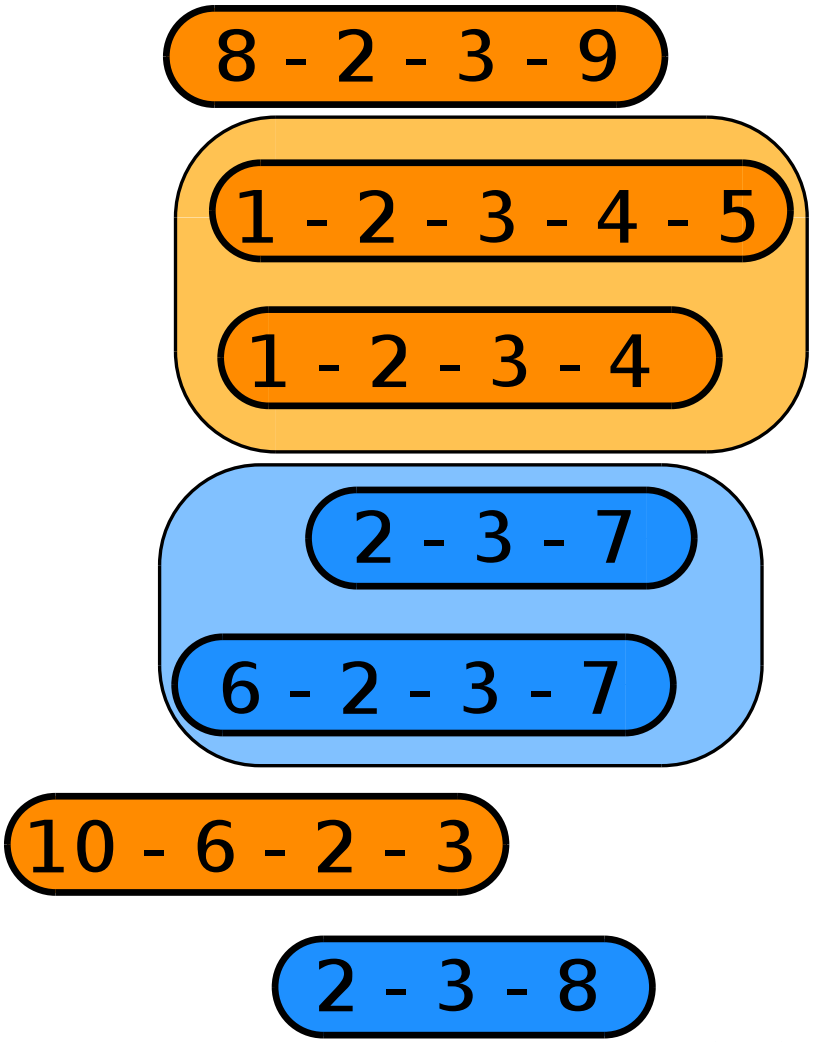
Path Clustering Example



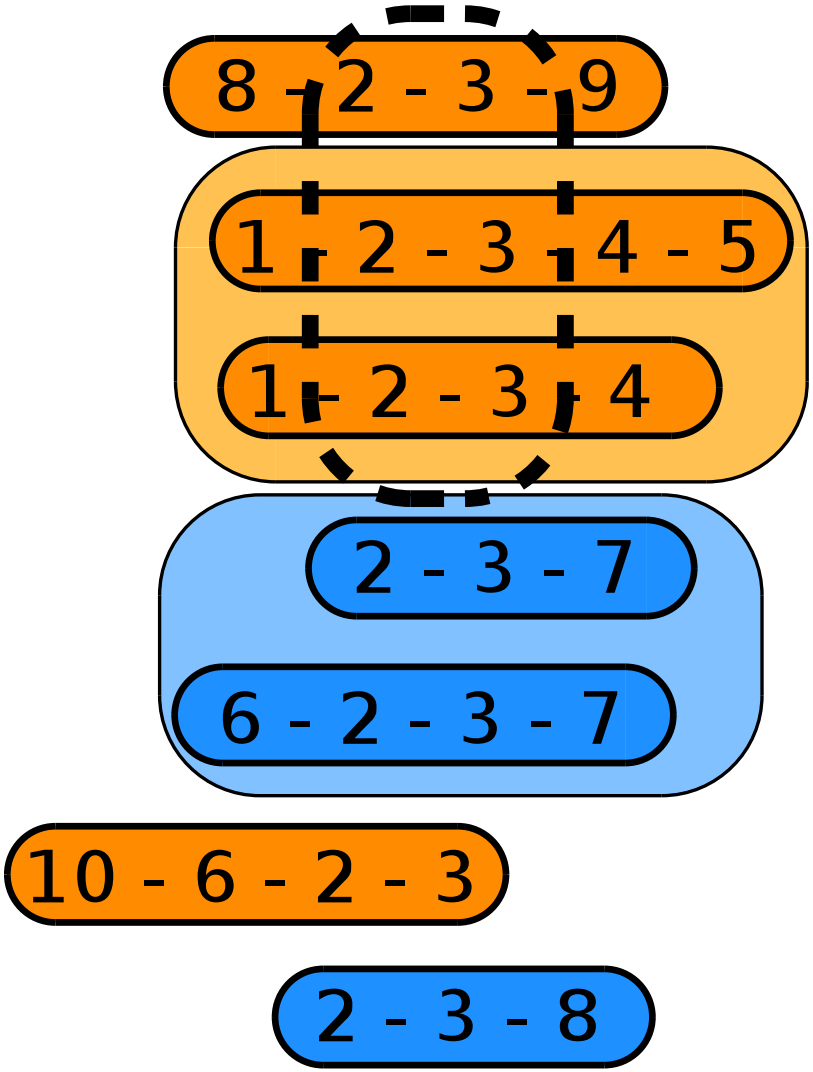
Path Clustering Example



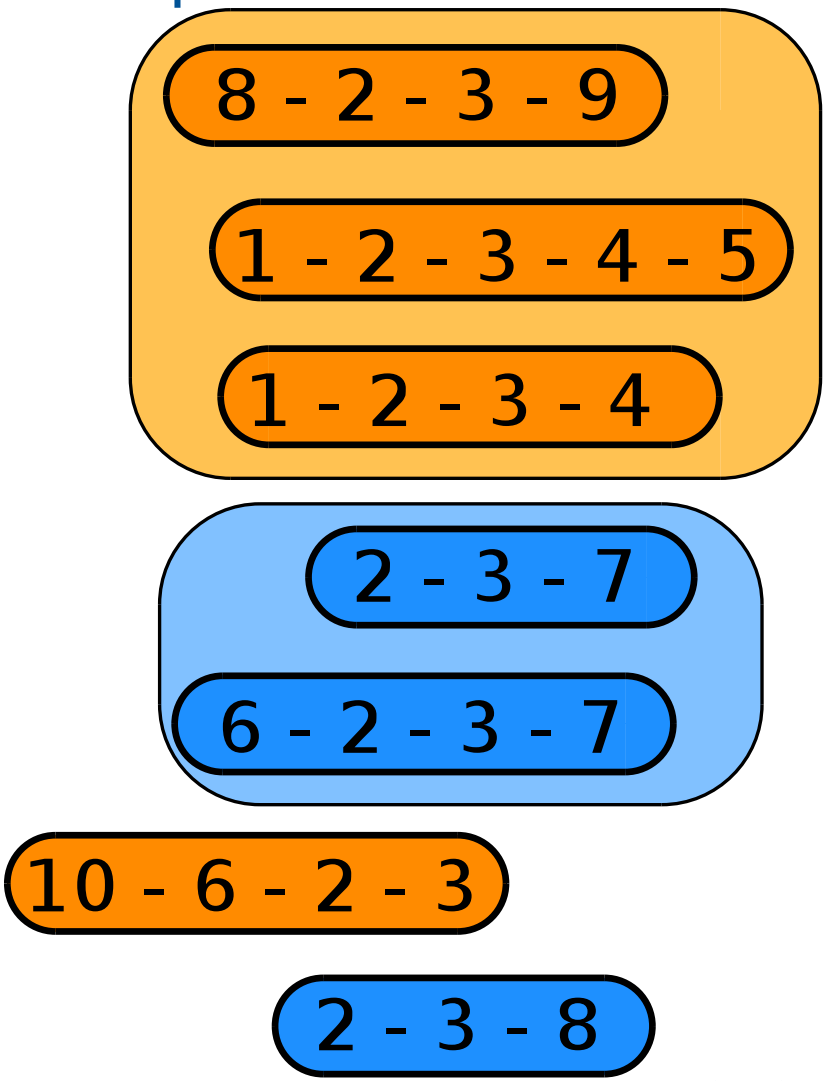
Path Clustering Example



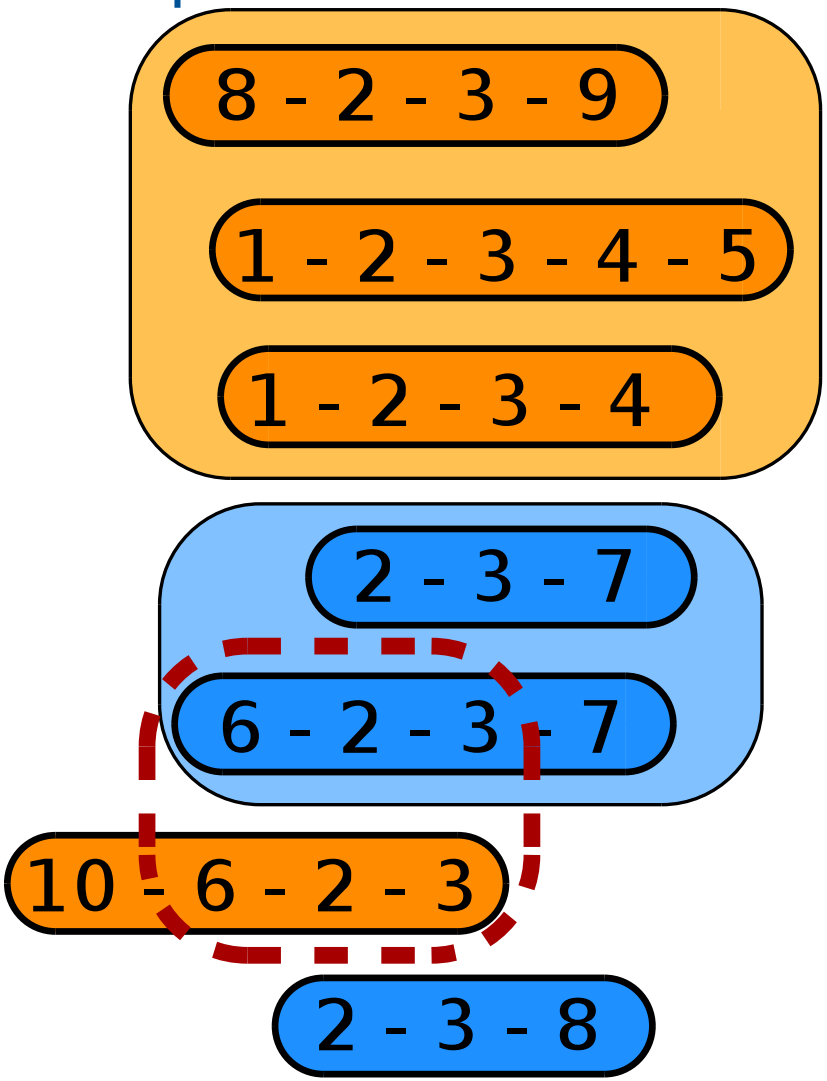
Path Clustering Example



Path Clustering Example



Path Clustering Example

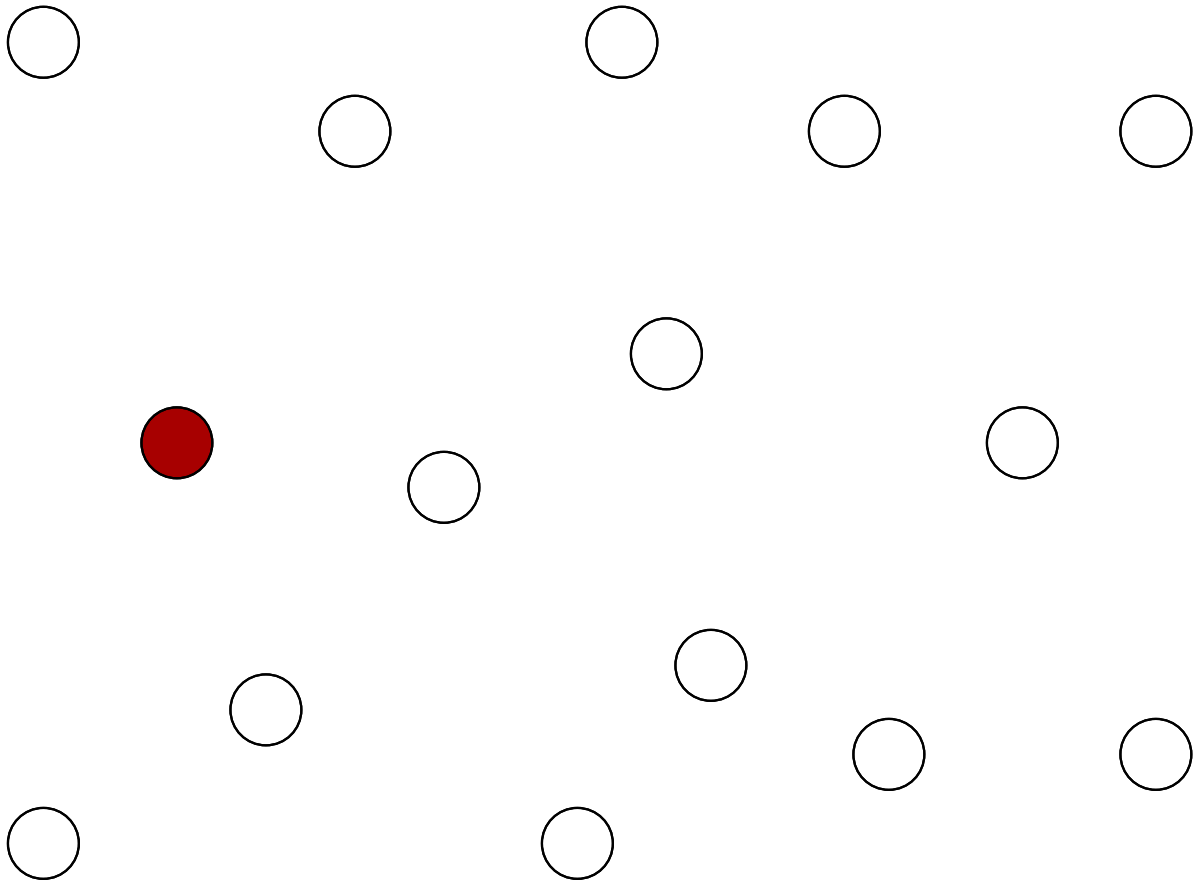


Decentralised Path Clustering

- To develop a scalable system the clustering approach must work without global knowledge.
- Cannot determine if a path is No-TIV without an exhaustive search of all paths
  - There is no No-TIV classification in the decentralised approach
- Two phases: **Measurement** and **Cluster Exchange**

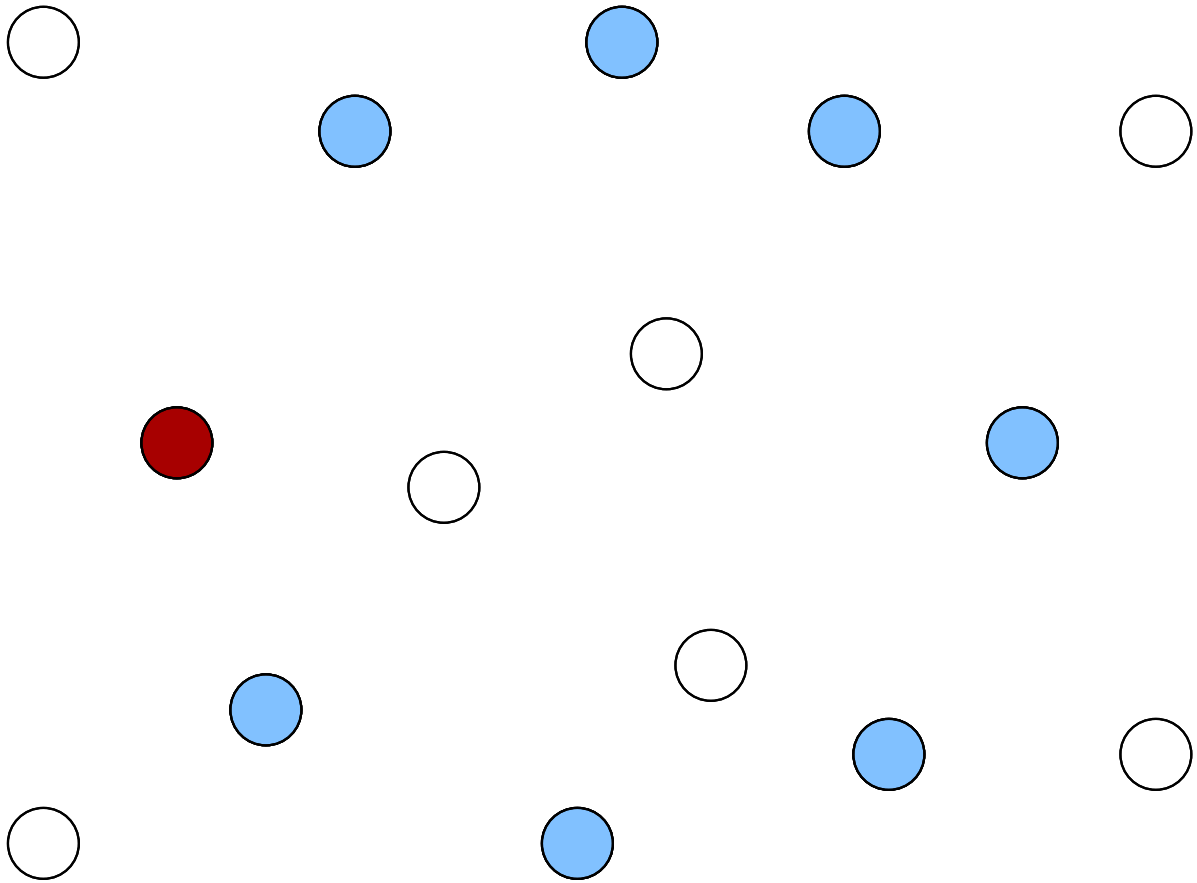
Decentralised Path Clustering: Measurement Phase

- Each node picks k of its n neighbours
- ...collaborates with those neighbours to measure complete end-to-end metrics and paths between themselves
- ... and identifies any resulting TIV paths which it can now cluster.



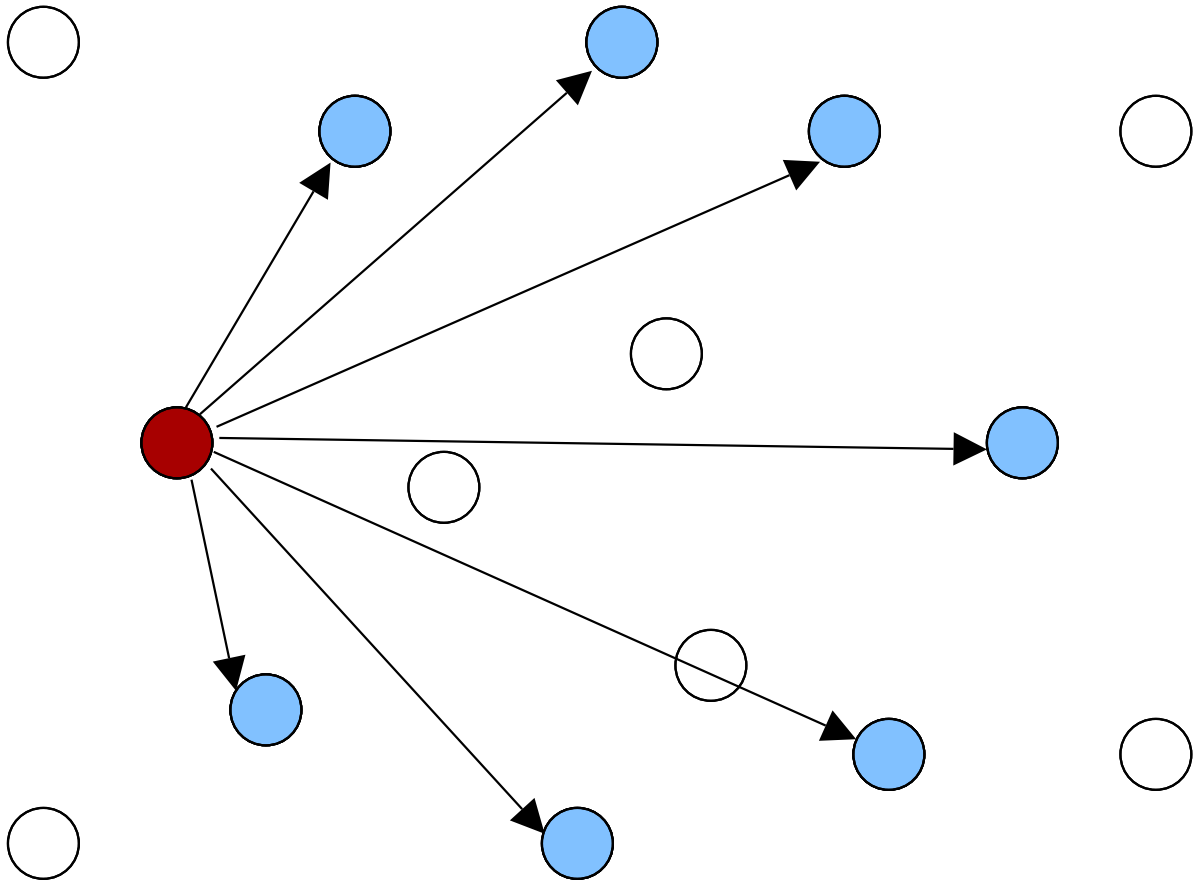
Decentralised Path Clustering: Measurement Phase

- **Each node picks k of its n neighbours**
- ...collaborates with those neighbours to measure complete end-to-end metrics and paths between themselves
- ... and identifies any resulting TIV paths which it can now cluster.



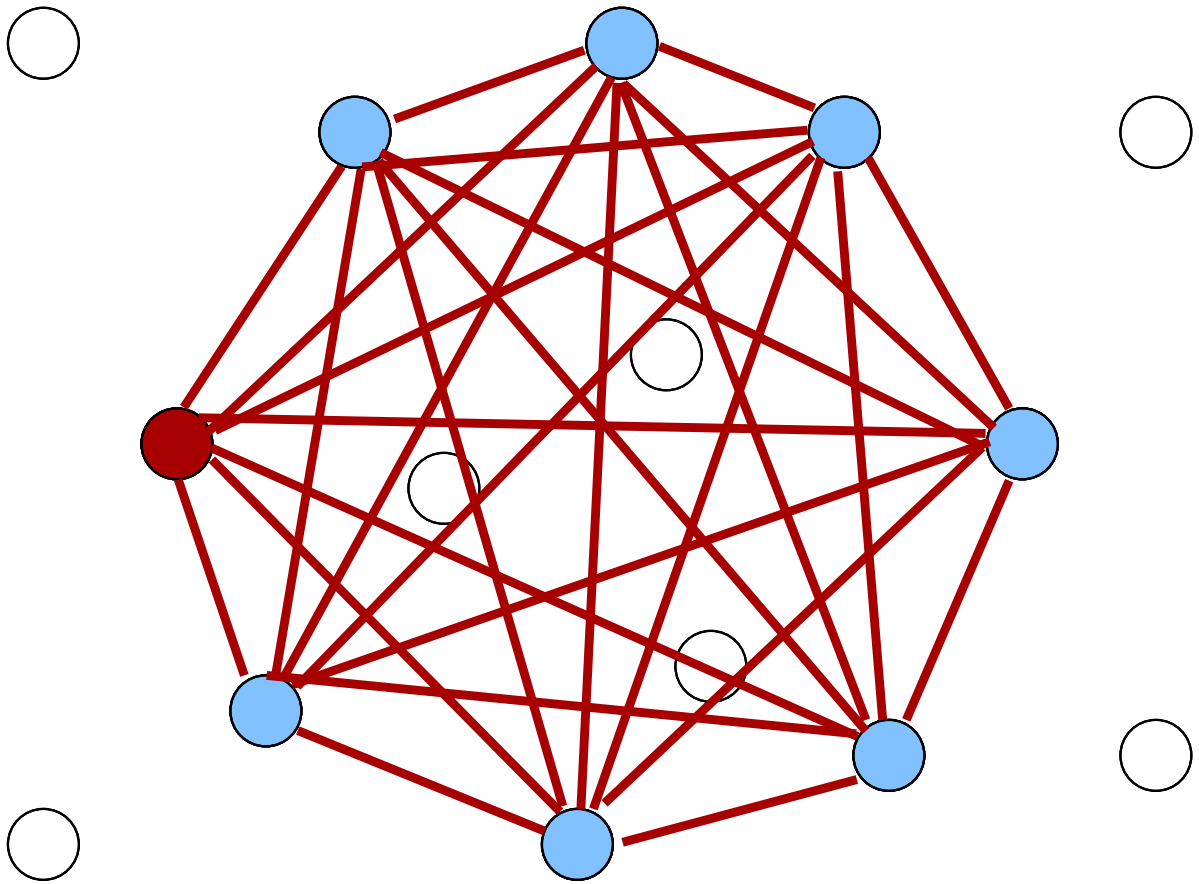
Decentralised Path Clustering: Measurement Phase

- Each node picks k of its n neighbours
- ...**collaborates** with those neighbours to measure complete end-to-end metrics and paths between themselves
- ... and identifies any resulting TIV paths which it can now cluster.



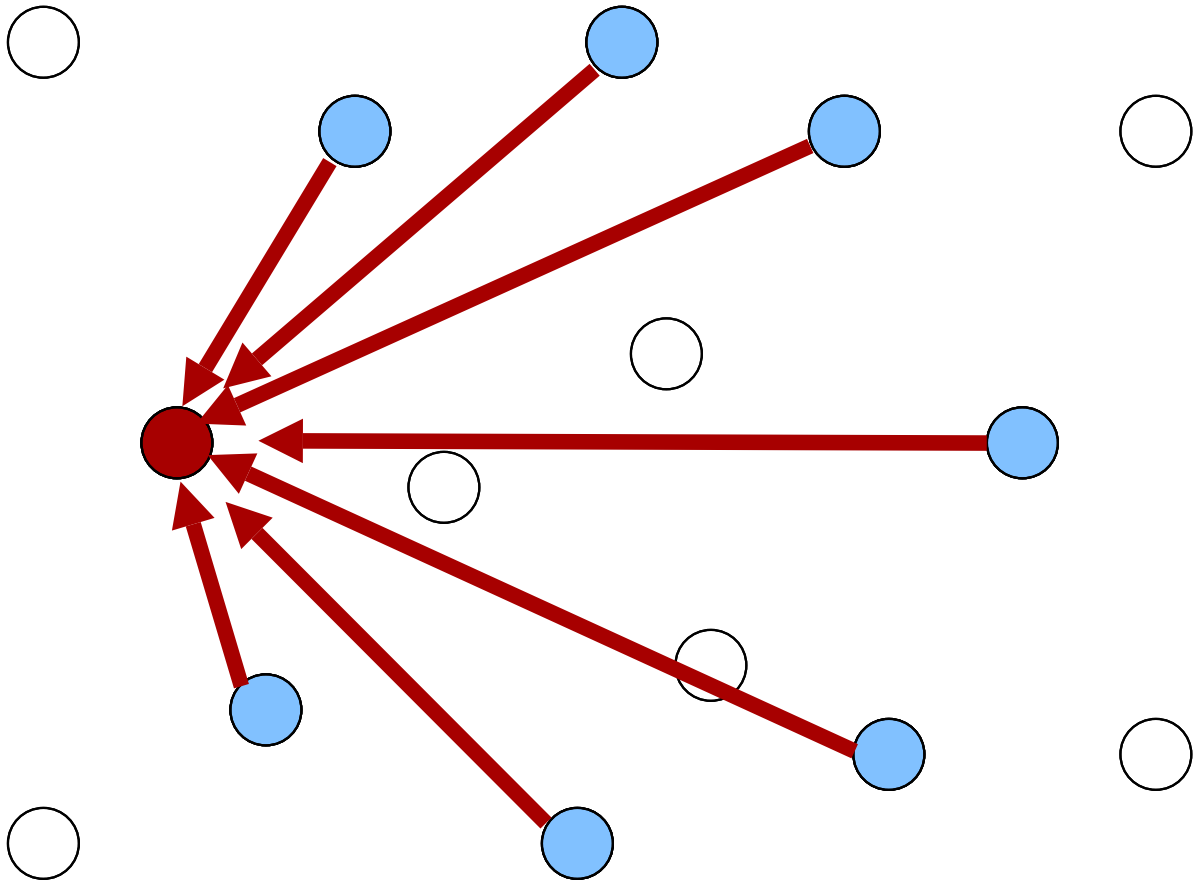
Decentralised Path Clustering: Measurement Phase

- Each node picks k of its n neighbours
- ...collaborates with those neighbours to **measure complete end-to-end metrics and paths between themselves**
- ... and identifies any resulting TIV paths which it can now cluster.



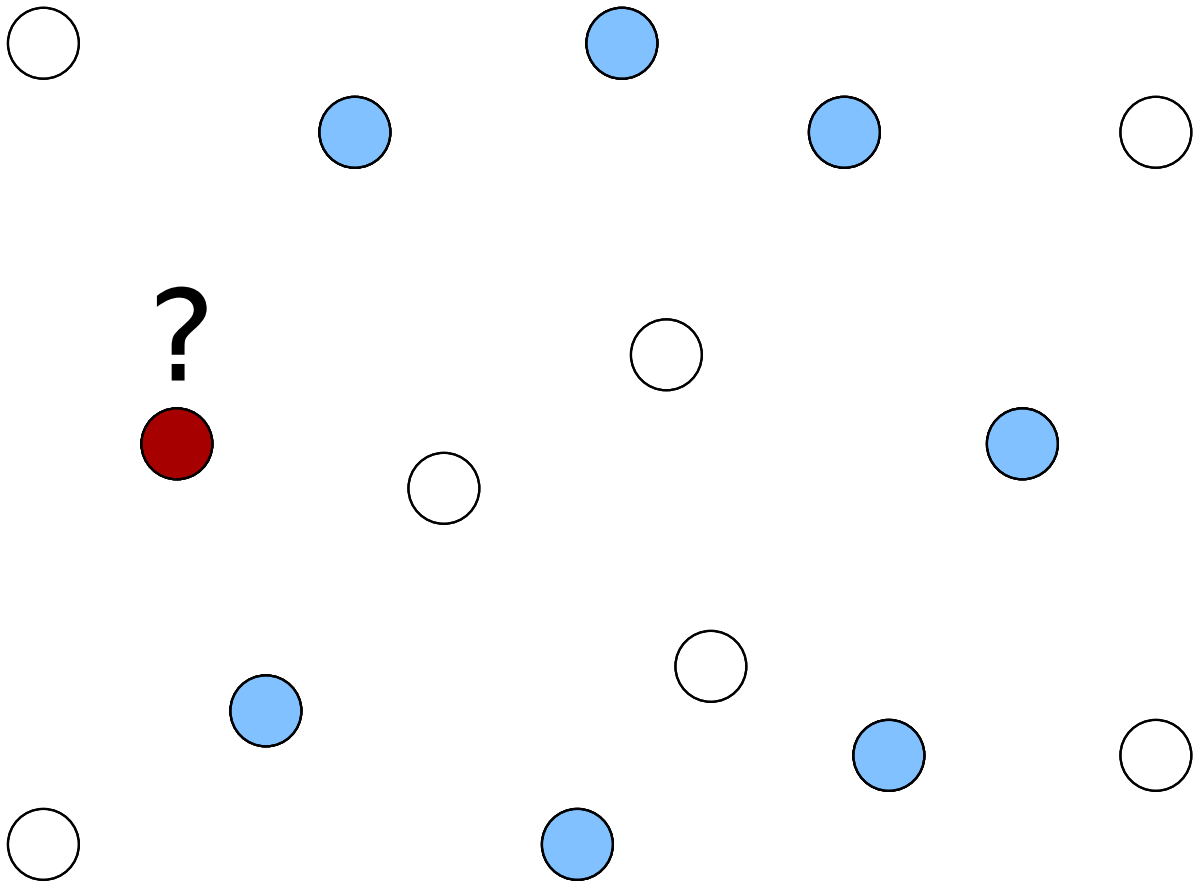
Decentralised Path Clustering: Measurement Phase

- Each node picks k of its n neighbours
- ...**collaborates with those neighbours to measure complete end-to-end metrics and paths between themselves**
- ... and identifies any resulting TIV paths which it can now cluster.



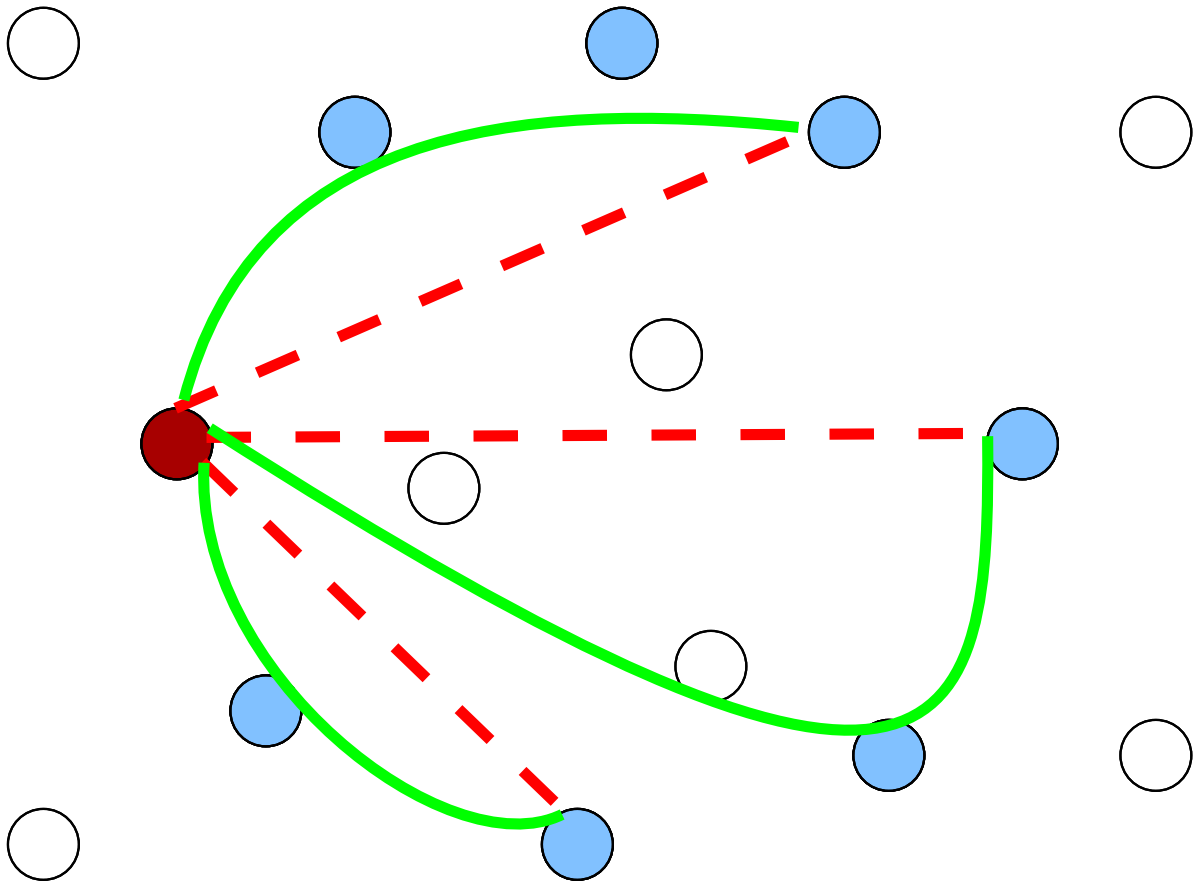
Decentralised Path Clustering: Measurement Phase

- Each node picks k of its n neighbours
- ...collaborates with those neighbours to measure complete end-to-end metrics and paths between themselves
- ... **and identifies any resulting TIV paths** which it can now cluster.



Decentralised Path Clustering: Measurement Phase

- Each node picks k of its n neighbours
- ...collaborates with those neighbours to measure complete end-to-end metrics and paths between themselves
- ... **and identifies any resulting TIV paths** which it can now cluster.



Decentralised Path Clustering: Cluster Exchange Phase

- Each node now has a small set of clusters based on local measurements
- At this point a number of rounds of cluster exchange occur:
- **Cluster Exchange:**
  1. Nodes choose a random neighbour and request its cluster set
  2. Upon receiving a new cluster set, nodes merge it with their own and rerun the clustering algorithm.

Centralised Evaluation

- We measured a 179 node all-to-all latency and AS path dataset on PlanetLab
  - 79% of direct paths formed TIVs
- Performed path clustering on all paths
  - Formed 5257 clusters (52.4% TIV) from 20614 paths
- **Aiming to show the upper baseline in cluster classification and validate the clustering approach**, we re-classified all paths against the resulting cluster set:

	Estimate	
	Correct	Incorrect
Have Detours (TIV paths)	94.3%	5.7%
No Detours (No-TIV paths)	83.1%	16.9%
Overall	90.6%	9.4%

Centralised Graph

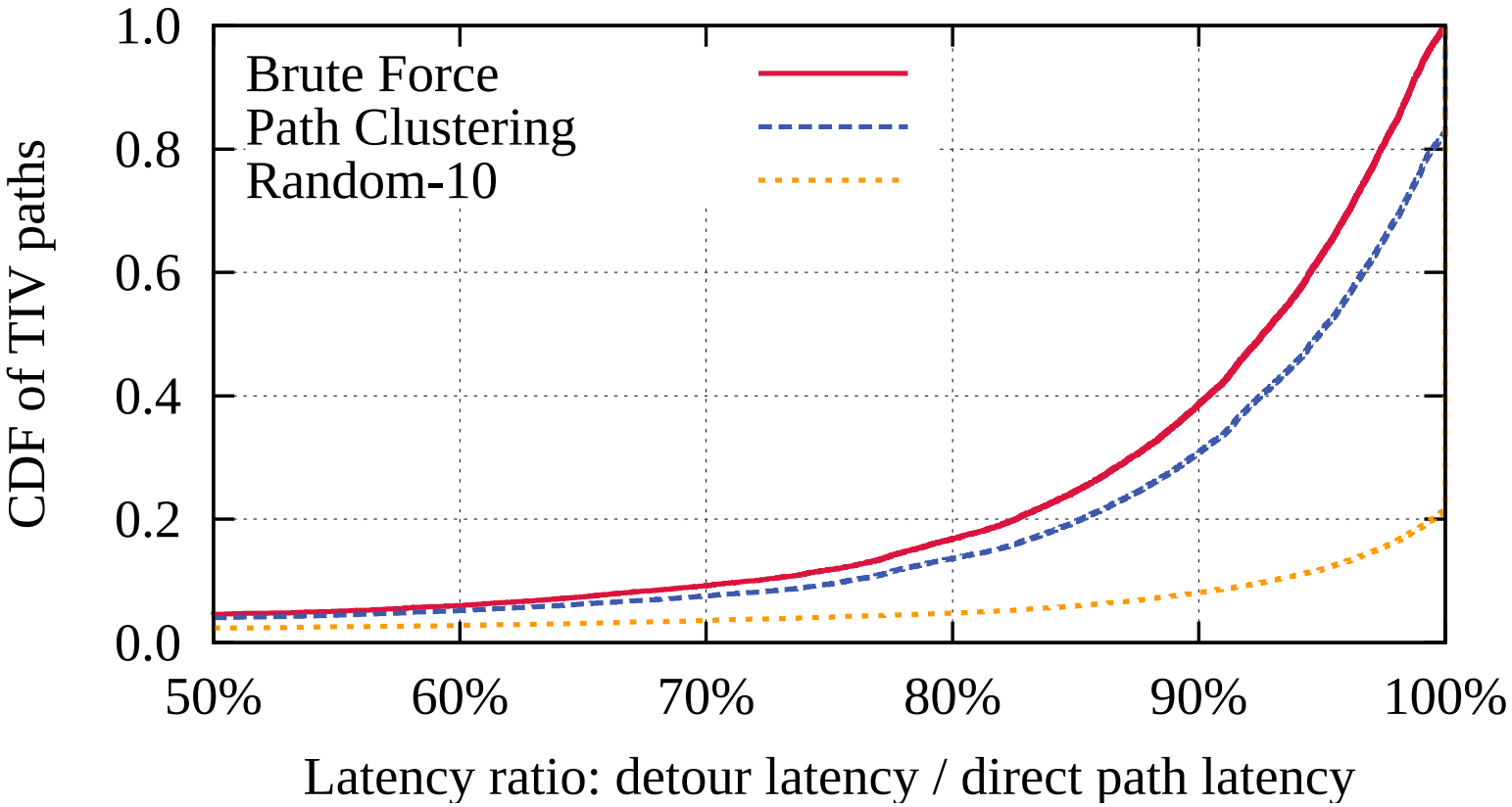
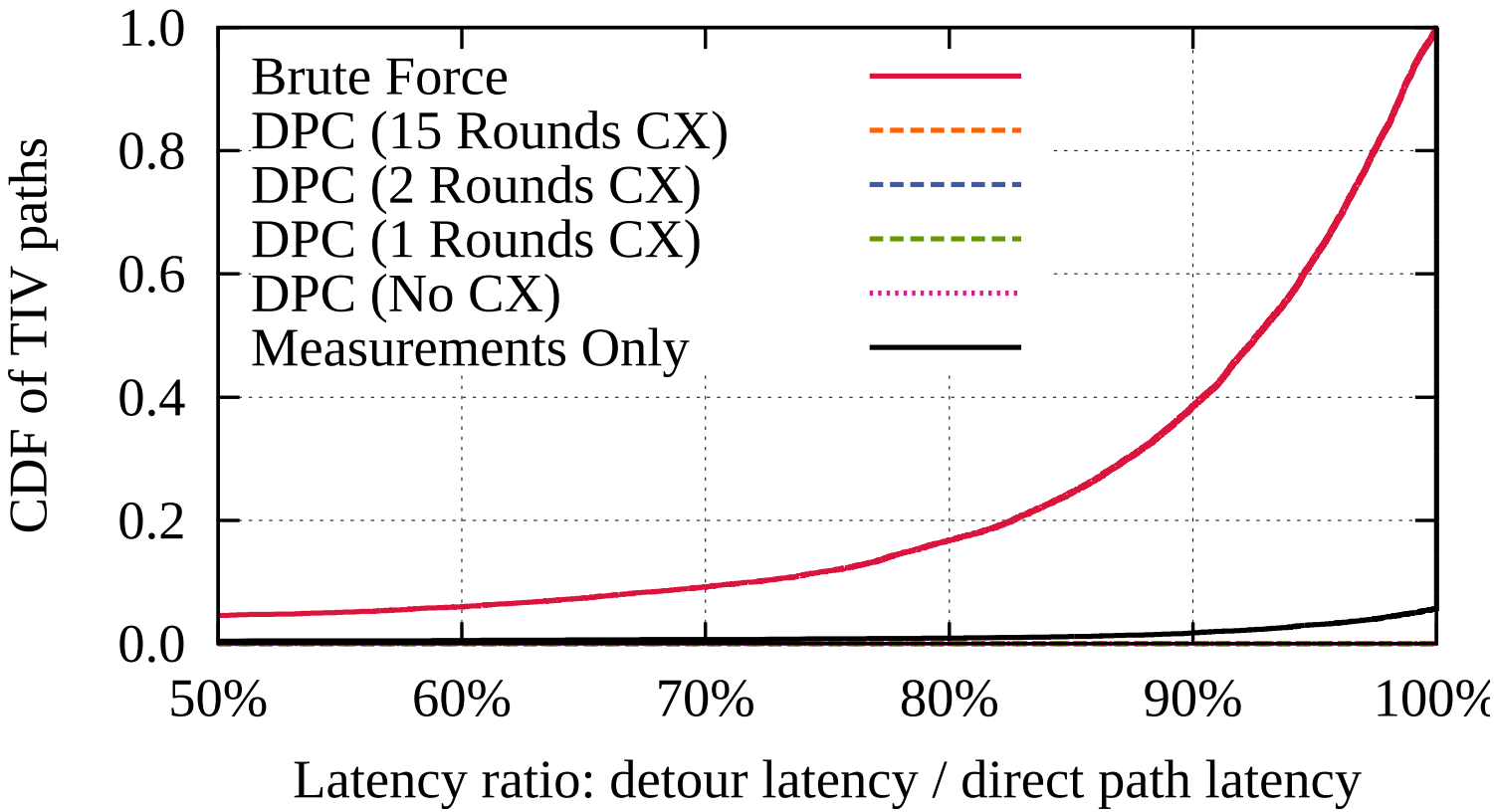


Figure: Top two cluster detours tested. Least lossy clustering

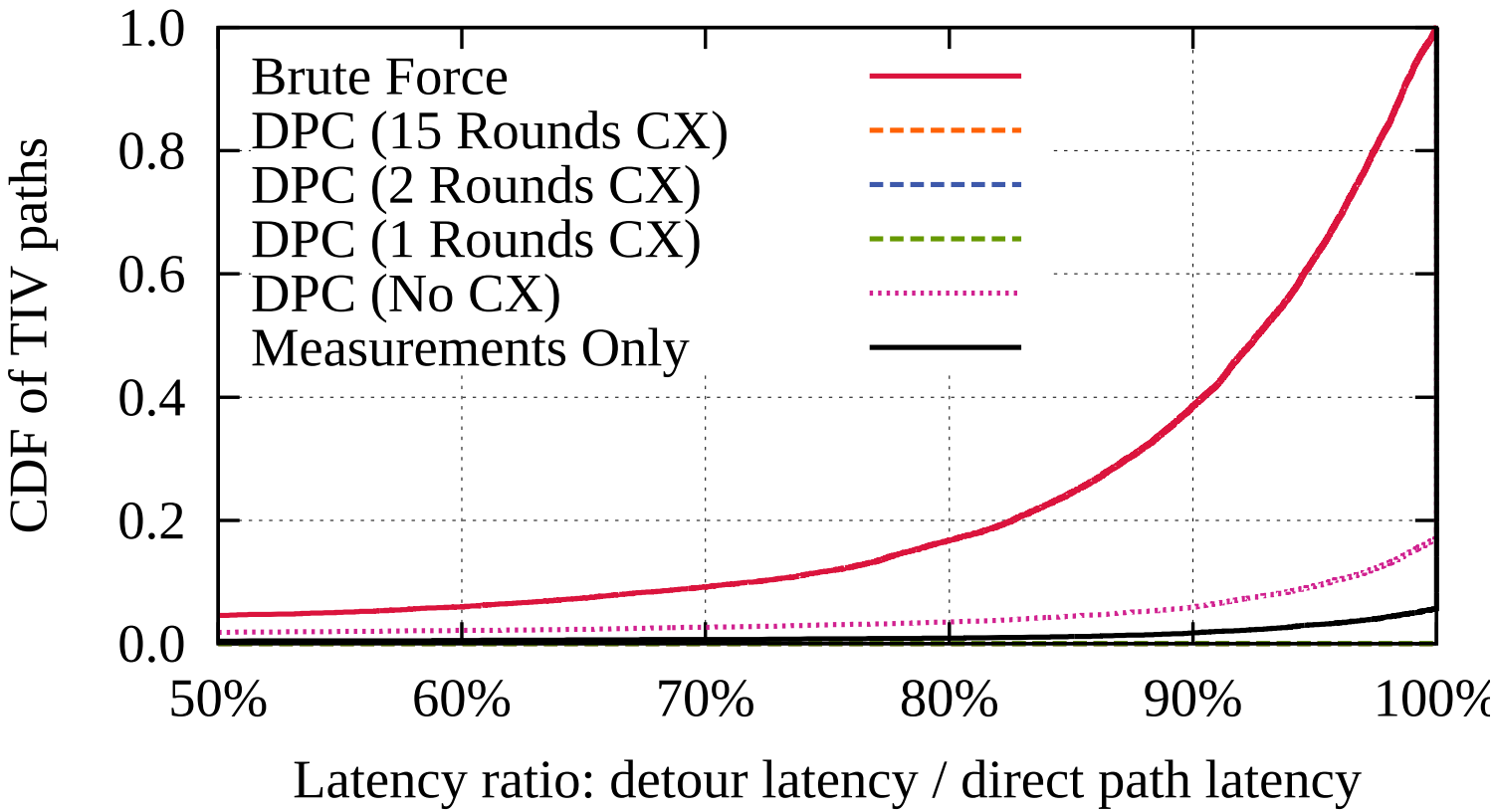
Decentralised Evaluation

- **Aiming to show the effectiveness of path clustering in identifying detours when we only have access to a small subset of measurements**
- Again we simulated based on the 179 node dataset
- Chose $k = 17$ (number of neighbours for each node to measure between)
- Each node found on average 59 TIV paths in its measurement subset, which formed on average 32 clusters.

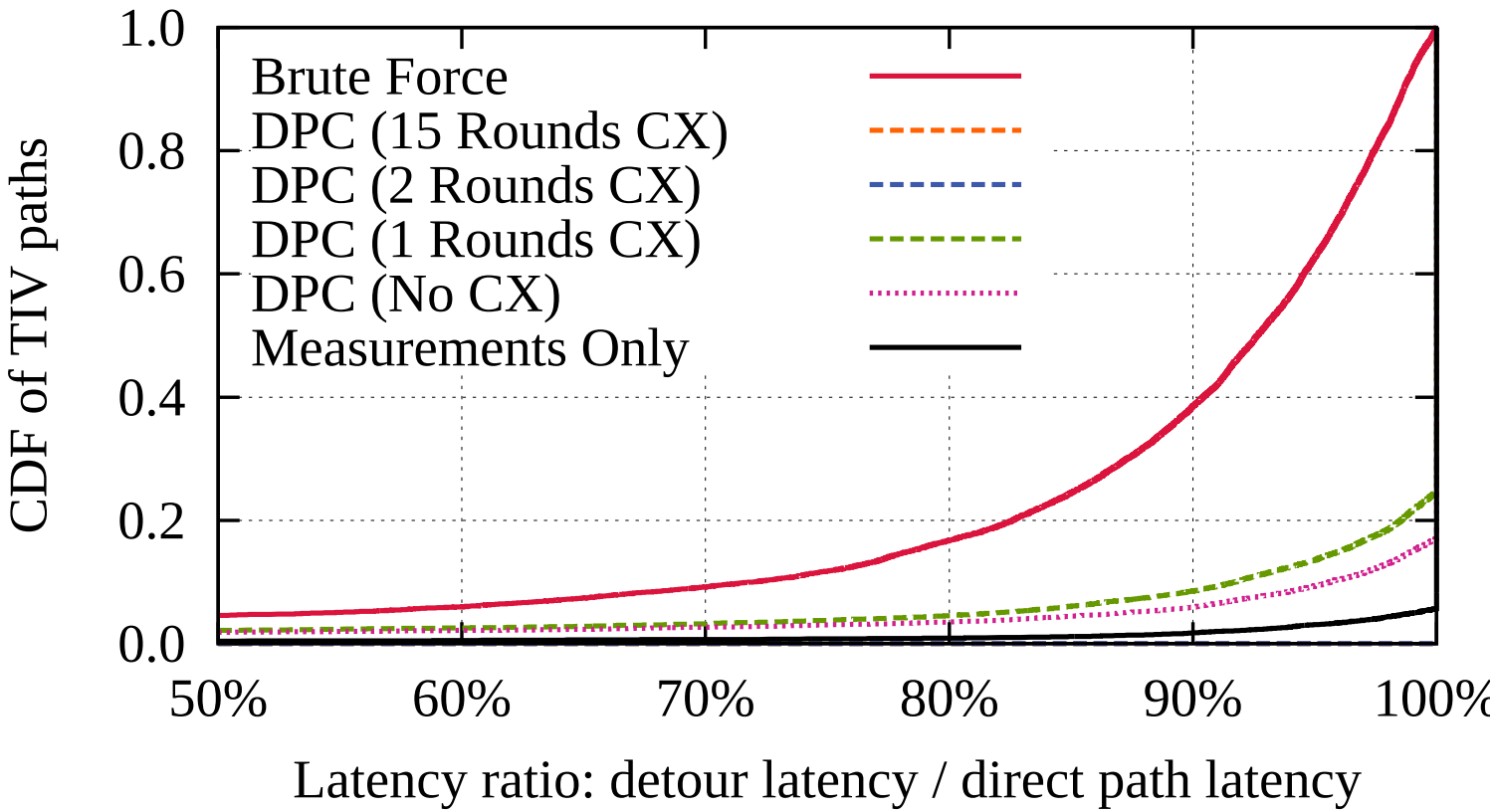
Decentralised Graph



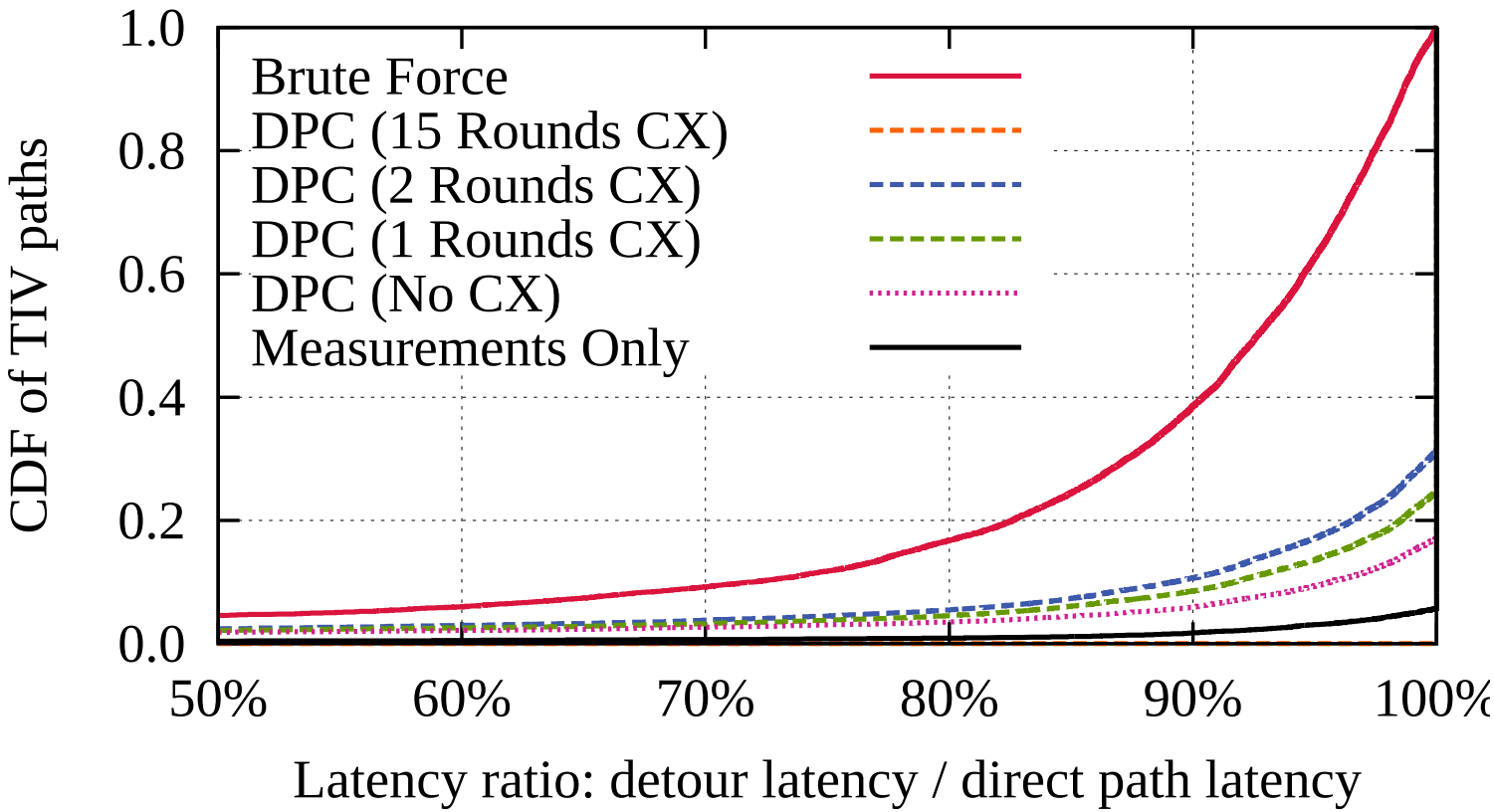
Decentralised Graph



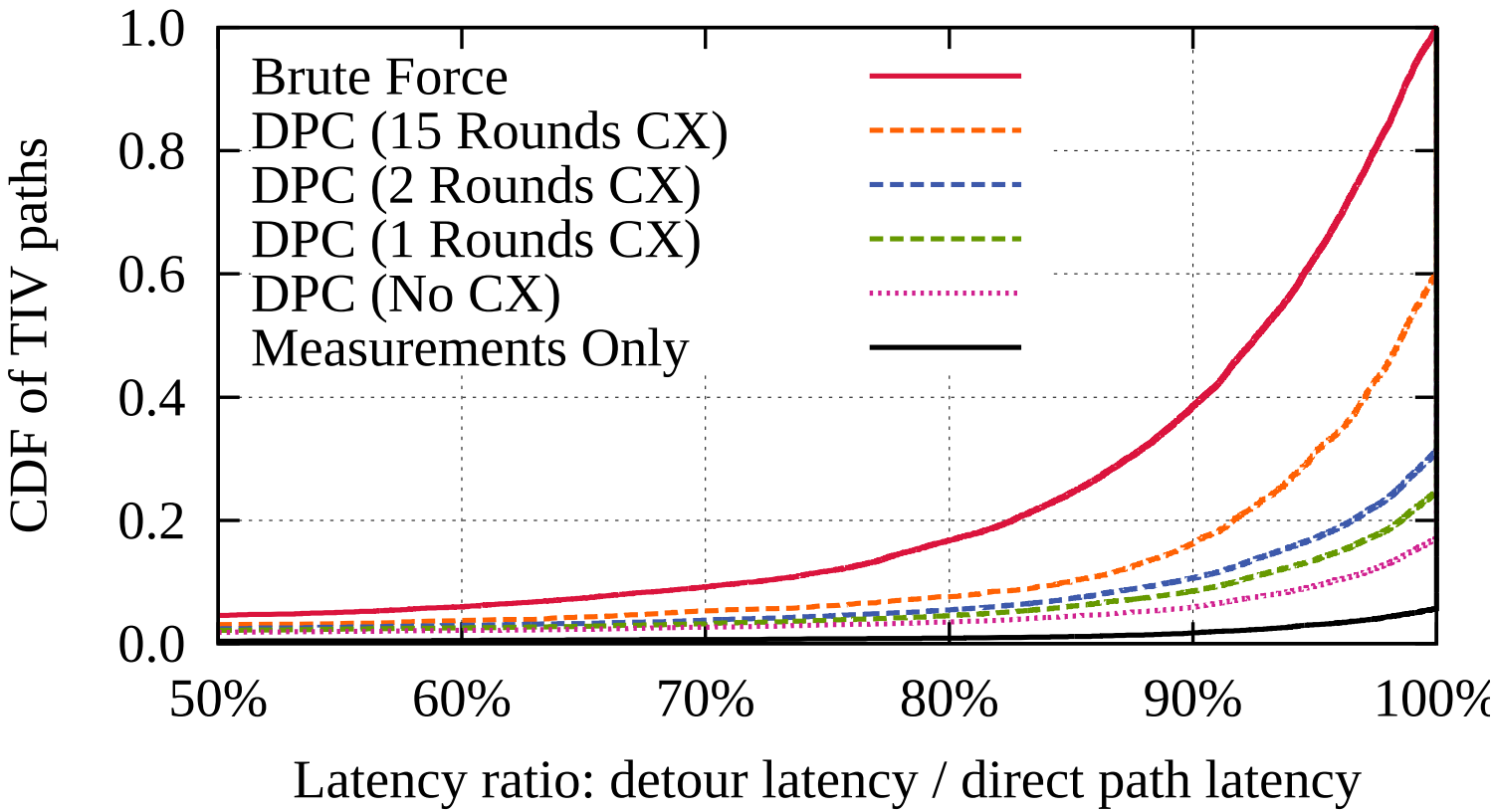
Decentralised Graph



Decentralised Graph



Decentralised Graph



Conclusions

- **AS path similarity is an effective indicator of transferability of detours between nodes**
- **AS path clustering can enable detours to be located without all-to-all end-to-end measurements**

Future directions

- Collecting datasets for other metrics
- Cluster compression
- Smarter heuristics for selecting detours from clusters
  - Ranking by detour improvement instead of frequency of occurrence
  - Randomising selection to improve load distribution
- Currently developing PlanetLab based routing plane to enable real end-user applications to exploit various forms of detour routing

Future directions

- Collecting datasets for other metrics
- Cluster compression
- Smarter heuristics for selecting detours from clusters
  - Ranking by detour improvement instead of frequency of occurrence
  - Randomising selection to improve load distribution
- Currently developing PlanetLab based routing plane to enable real end-user applications to exploit various forms of detour routing

Questions?

Thom Haddow — thaddow@doc.ic.ac.uk