

Bloom Filter based Inter-domain Name Resolution: A Feasibility Study

Konstantinos V. Katsaros, Wei Koong Chai and George Pavlou
University College London, UK



Outline

- Inter-domain name resolution in ICN
- Scalability concerns
- Bloom filter based name resolution
- Evaluation framework
- Results
- Conclusions and Future Work

Inter-domain name resolution in ICN

- Name resolution: taking forwarding decisions based on names
- Inter-domain level → Enormous size of the namespace
 - More than a trillion (10^{12}) unique web pages (Google)
 - More than 50 billion (10^9) IoT devices expected (Cisco)
 - Other estimations for 10^{16} Information Objects (IOs)
 - Exact size subject to naming granularity i.e., hierarchical vs. flat
- Concerns about scalability
 - Memory: maintain state in RAM for low latency
 - Processing: lookup overheads
 - Bandwidth: propagate state

Inter-domain name resolution in ICN

Lookup-by-name approaches

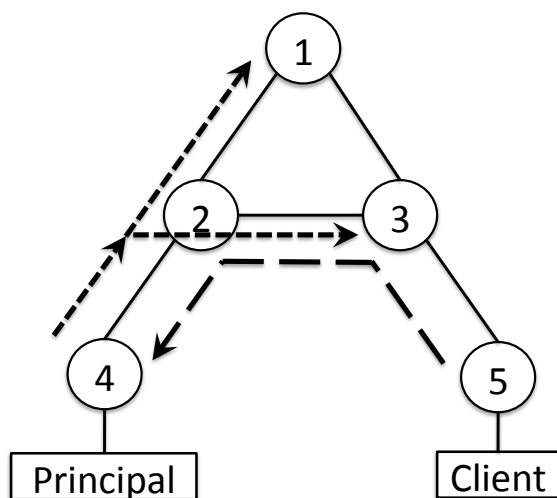
- Distributed directory service
 - Looking up forwarding / location information
- Usually based on Distributed Hash Tables (DHTs)
 - ✓ Perfect load balancing
 - ✗ Stretched name resolution paths
 - ✗ Routing policy violations
 - ✗ Limited control over state placement

K.V. Katsaros, et al., "On Inter-domain Name Resolution for Information-Centric Networks," IFIP-TC6 Networking, 2012

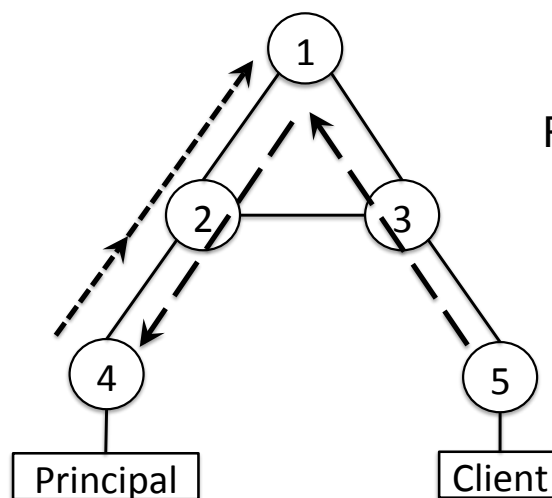
Inter-domain name resolution in ICN

Route-by-name approaches

- Name resolution state leads to content
- State replicated across the inter-domain topology following BGP routing
- ✓ Resolution paths follow the structure of the inter-domain topology



DONA (2007)



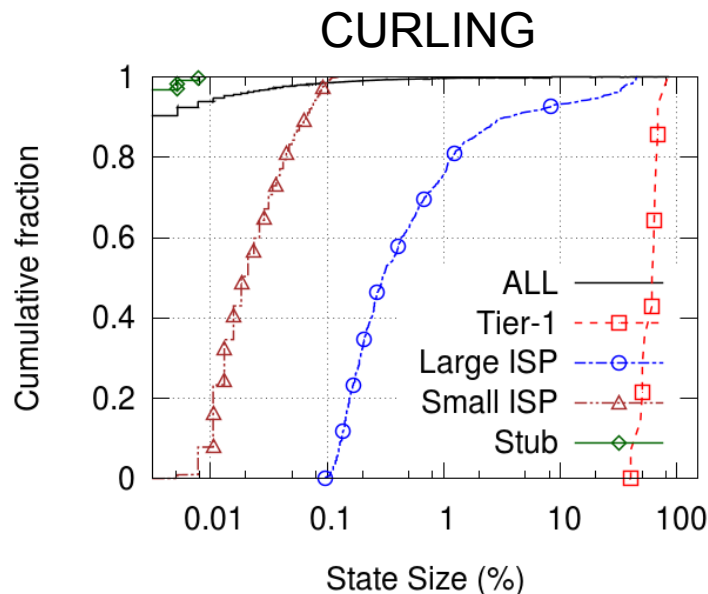
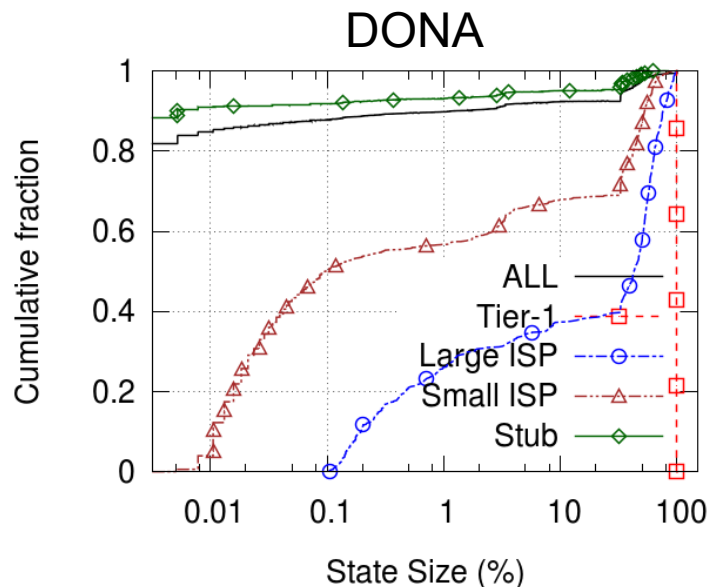
CURLING (2011)

REGISTRATION

FIND

Inter-domain name resolution in ICN

Route-by-name approaches



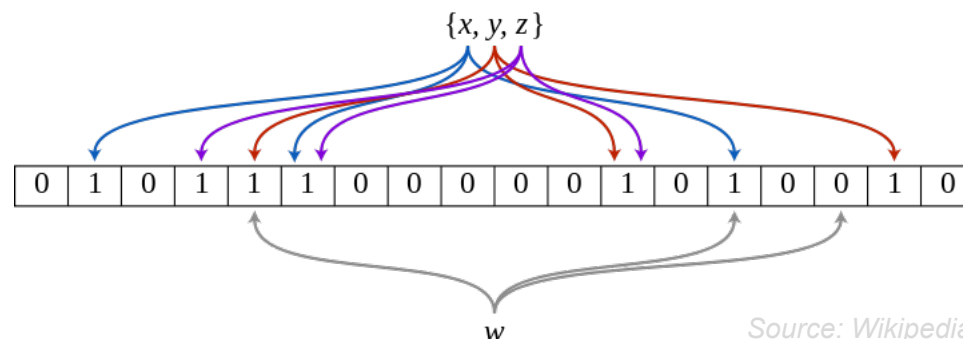
- ✗ State heavily replicated (DONA: x1702.64, CURLING: x27.34)
- ✗ 420 TB of state for 10^{13} IOs at Tier-1 in DONA
- ✗ Highly skewed distribution of load across tiers

K.V. Katsaros et al., "On the Inter-domain Scalability of Route-by-Name Information-Centric Network Architectures," IFIP-TC6 Networking, May 2015

USING BLOOM FILTERS

- Hong et al. Bloom Filter-based Flat Name Resolution System for ICN. Internet-Draft draft-hong-icnrg-bloomfilterbased-name-resolution-03.txt, IETF Secretariat, Mar. 2015.
- H. Liu et al. A multi-level DHT routing framework with aggregation. In Proc. of the 2012 ACM SIGCOMM Workshop on Information-centric networking (ICN'12), pages 43–48. ACM, 2012.

Bloom Filters (BF)

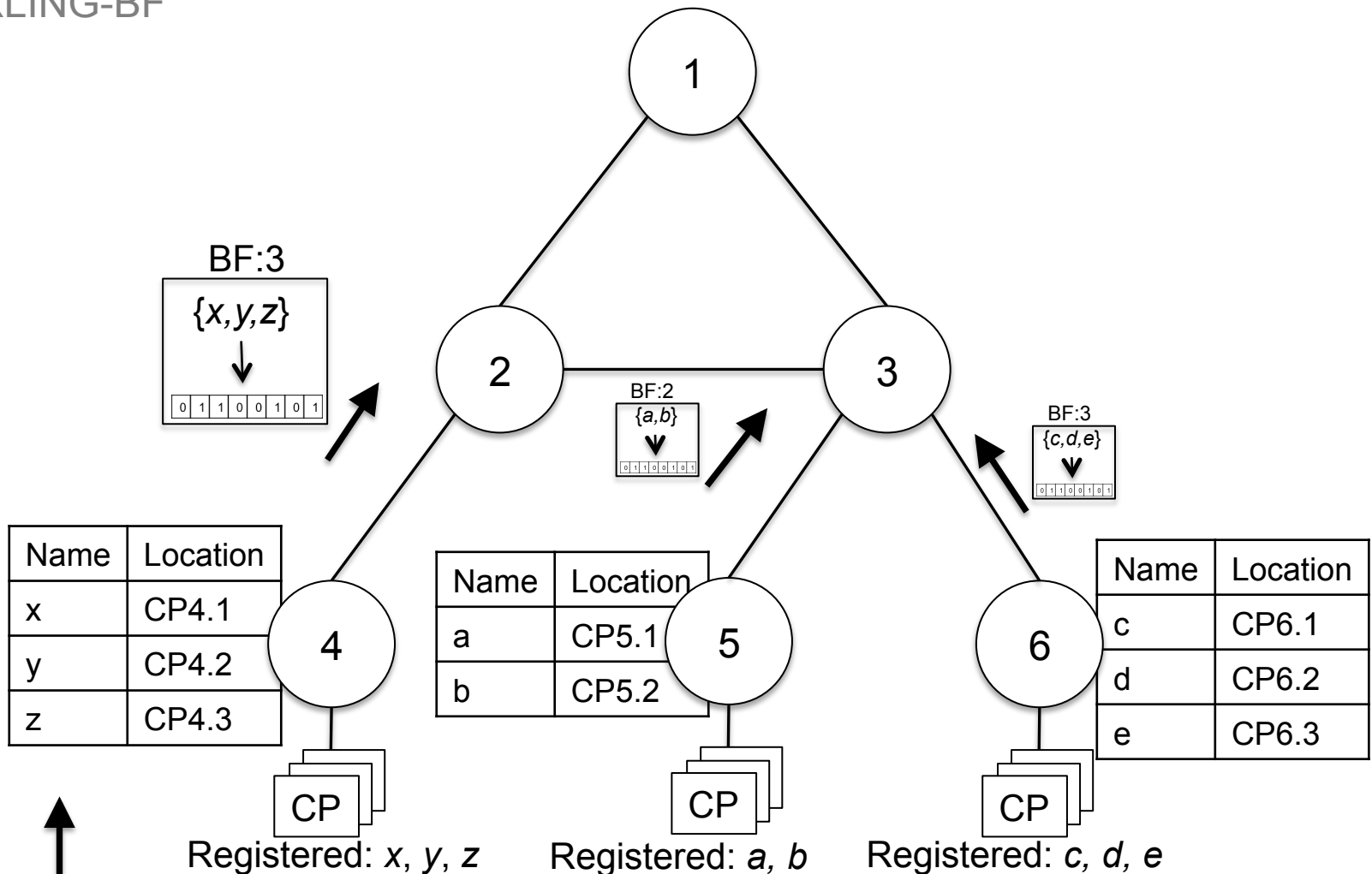


Source: Wikipedia

- Array of m bits
- k hash functions hash an element to one of the m positions
- **ADD**: hash element \rightarrow get k positions \rightarrow set to 1
- **QUERY**: hash element \rightarrow get k positions \rightarrow check if all set to 1
- **UNION**: bitwise OR
- False positive ratio (R): $R \approx (1 - e^{-kn/m})^k$
- Optimal number of hash functions: $k = \frac{m}{n} \ln 2$
- For R upper limit (R_{max}) and optimal k :
$$\frac{m}{n} = -\frac{\ln R_{max}}{(\ln 2)^2}$$
- For a given R_{max} and m we can calculate the *capacity* a BF (C_{BF})

Using Bloom Filters for Name Resolution

CURLING-BF

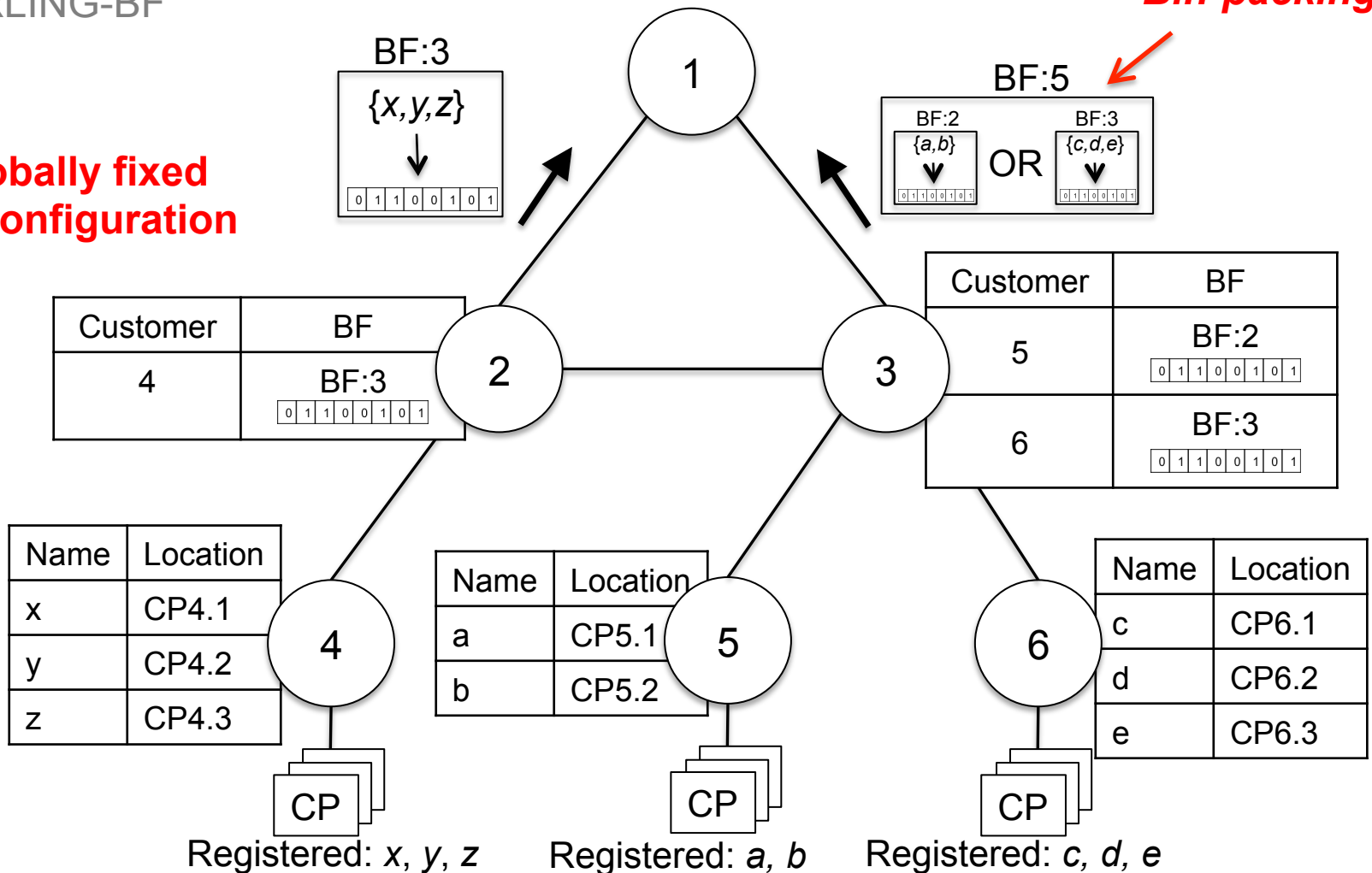


Using Bloom Filters for Name Resolution

CURLING-BF

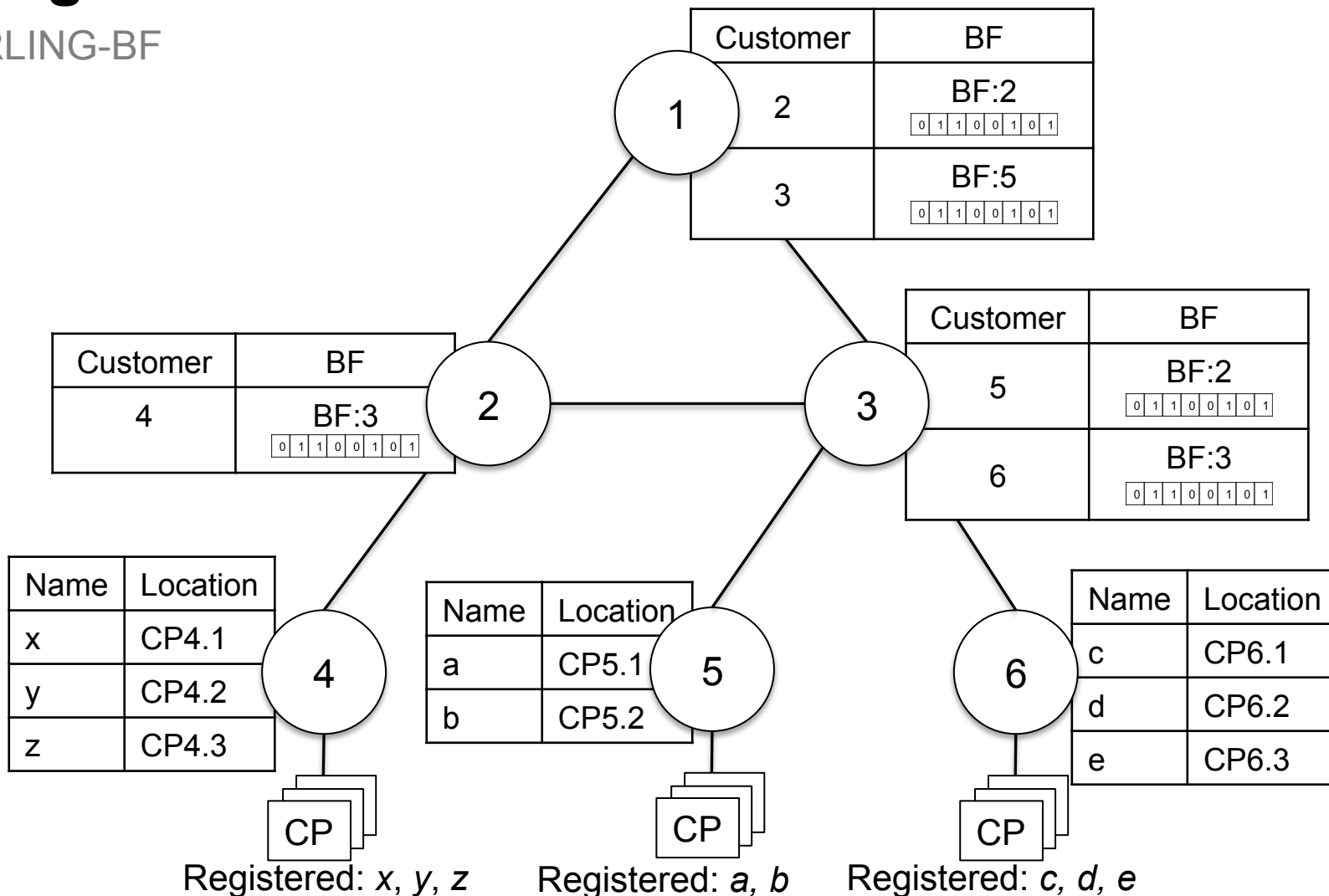
Globally fixed BF configuration

Bin-packing



Using Bloom Filters for Name Resolution

CURLING-BF



Configuring Bloom Filters for Name Resolution

- How to select m and C_{BF} ?
- **Primary objective:** limit false positives
- F : number of BFs at a node, s : number of registrations

$$F = \left\lceil \frac{s}{C_{BF}} \right\rceil$$

**Lower bound: overlooks BF table structure & assumes perfect bin-packing*

- Setting an upper limit for R at any node in the network (R_{max}^{Node})

$$R_{max}^{Node} = 1 - (1 - R_{max})^F$$

➤ Fixing R_{max}^{Node} for worst case i.e., tier-1 domains...

- Multiple conforming BF configurations i.e., $\langle m, C_{BF} \rangle$

$$m = -C_{BF} \frac{\ln(1 - (1 - R_{max}^{Node})^{\lceil \frac{s}{C_{BF}} \rceil})}{(\ln 2)^2}$$

BF Configuration Tradeoff: metrics

- Memory Requirements

$$M^x = b_{NonBF} \cdot s_l + b_{BF} \cdot F, \quad x \in \{CURLING-BF, DONA-BF\}$$

$$M^x = b_{NonBF} \cdot s, \quad x \in \{CURLING, DONA\}$$

$$b_{NonBF} = 42, \quad b_{BF} = (m + \log_2 C_{BF})/8$$

- Processing Overheads

$$LO^x = \alpha \cdot \log(s) + c, \quad x \in \{CURLING, DONA\}$$

$$LO^x = F \cdot P_N^F + \frac{P_P}{P_N} \cdot \sum_{i=0}^F i P_N^i, \quad x \in \{CURLING-BF, DONA-BF\}$$

$$p = \frac{s}{S}, \quad P_{TP} = \frac{p}{F}, \quad P_P = (1 - P_{TP})R_{max}^{FP} + P_{TP}, \quad P_N = 1 - P_P$$

Resource requirements depend on the number of BFs

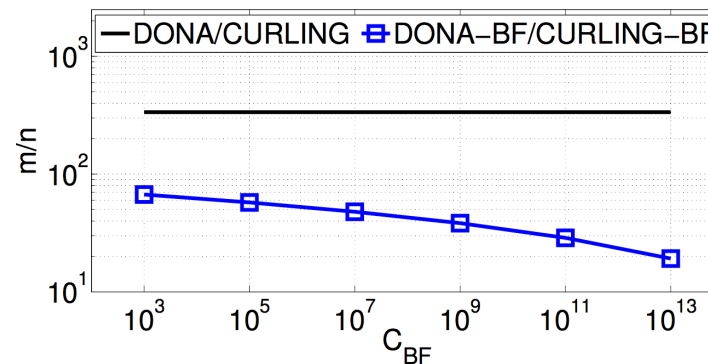
BF Configuration Tradeoff

$$C_{BF} \gg s$$

- Sparse BFs → resource waste: memory and bandwidth

$$C_{BF} \ll s$$

- Multitude of BFs → increased processing overheads
 - Increased number of bits-per-elements to support R_{max}^{Node}

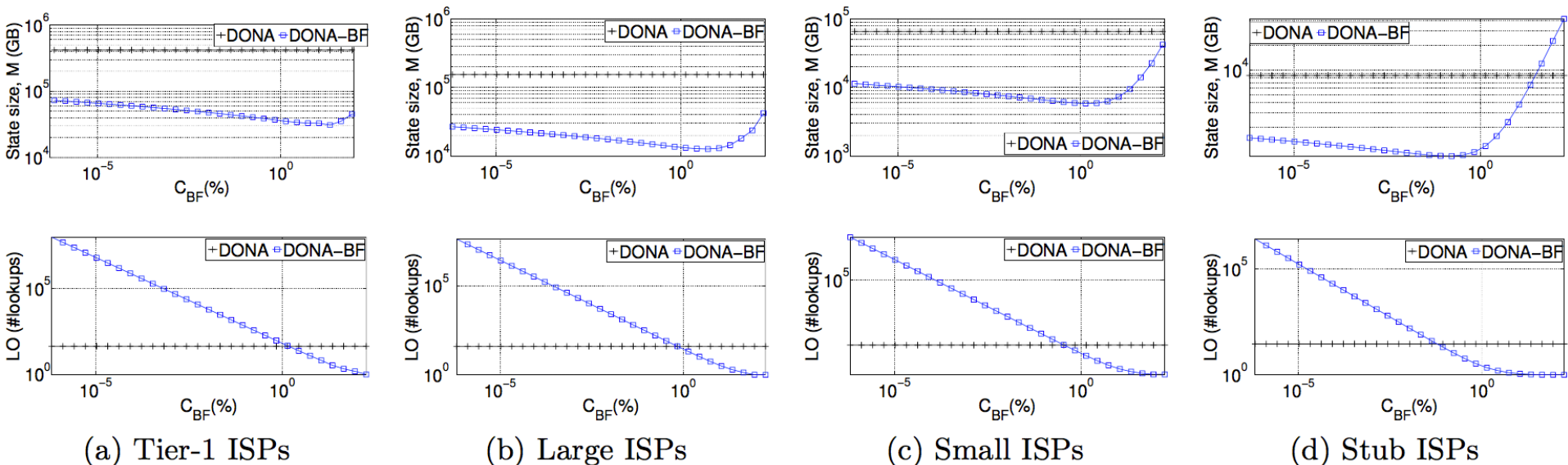


$$C_{BF} = s$$

- Ideal, but state distribution is heavily skewed i.e., no single s value ...

Empirical Observations (CAIDA trace set)

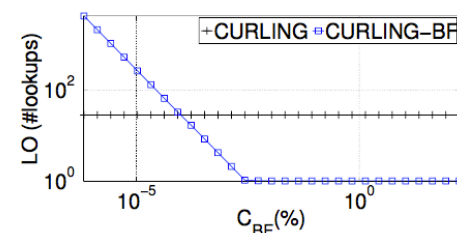
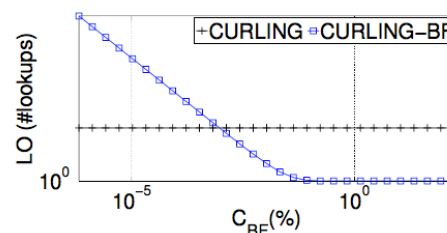
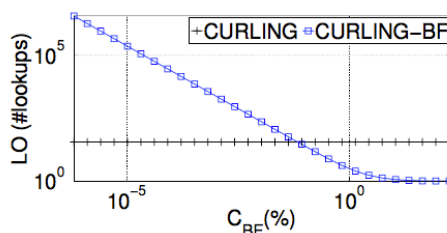
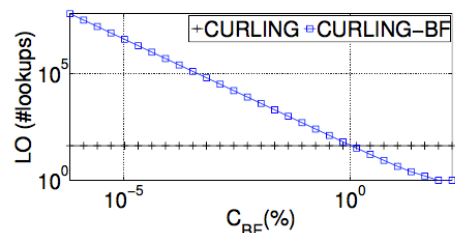
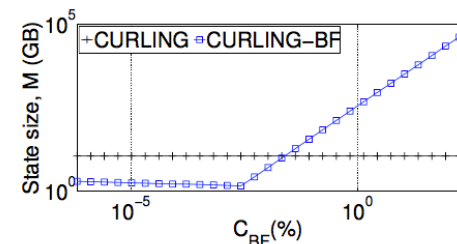
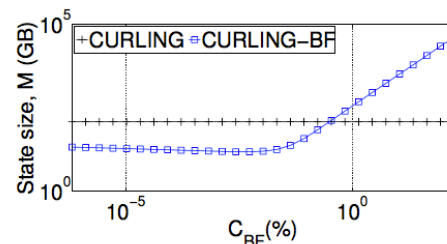
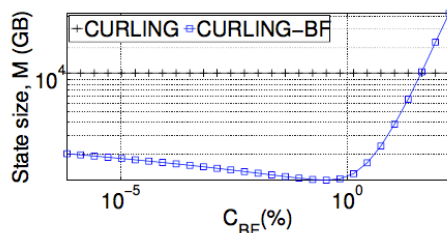
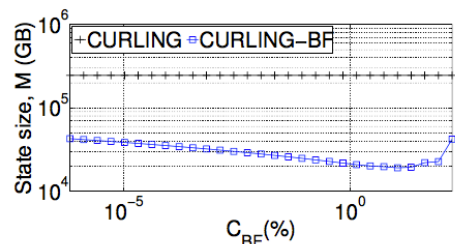
DONA/DONA-BF



- **For large C_{BF} :** e.g. 10^{13} IOs $\rightarrow m = 23.96$ TB for a single BF!
 - Practical RAM limitations
 - No incentives for Stub ASes to use BFs (memory)
- **For small C_{BF} :** e.g., 10^5 IOs $\rightarrow m = 718.88$ KB $\rightarrow 10^8$ BF lookups at tier-1
 - No incentives for Tier-1, Large ISPs to use BFs (processing)

Empirical Observations (CAIDA trace set)

CURLING/CURLING-BF



(a) Tier-1 ISPs

(b) Large ISPs

(c) Small ISPs

(d) Stub ISPs

- **No single BF configuration can yield both lower memory and processing resource requirements for all ASes**

Empirical Observations (CAIDA trace set)

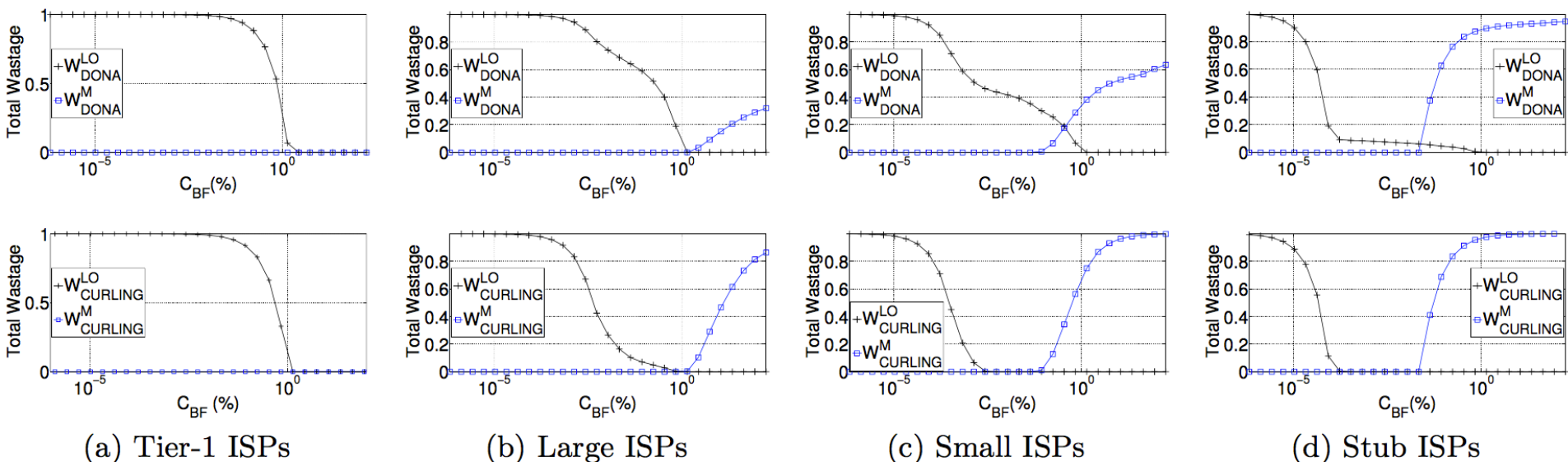
- Are there BF configurations leading to *some* incentives for *all* ASes?
- *Resource wastage*

$$w_{i,x-BF}^y = \max\left\{0, 1 - \frac{y_i^x}{y_i^{x-BF}}\right\}$$

$$W^y = \frac{\sum_{i=1}^N w_{i,x-BF}^y}{N}$$

$$y \in \{M, LO\}, x \in \{DONA, CURLING\}, i \in [1, \dots, N]$$

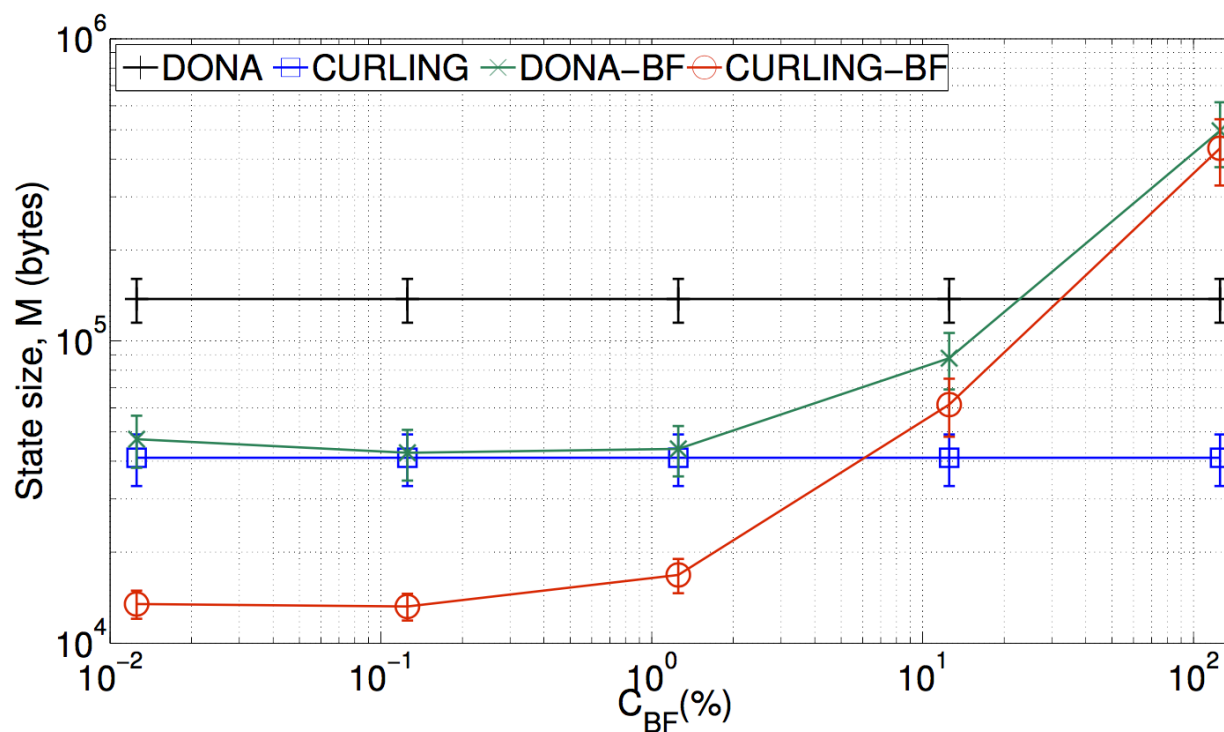
Empirical Observations (CAIDA trace set)



- Stub networks (vast majority of ASes)
 - Low resource wastage range: ($C_{BF} = 2^{23}$, $m = 50.63\text{MB}$) to ($C_{BF} = 2^{32}$, $m = 19\text{GB}$)
 - But substantial processing overheads for Tier-1/Large ISPs at this range
- **No single BF configuration can achieve a *good* compromise**

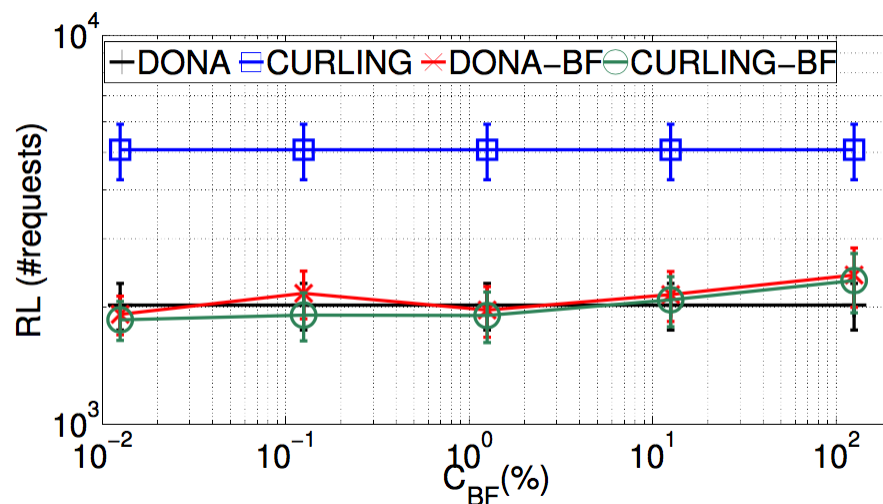
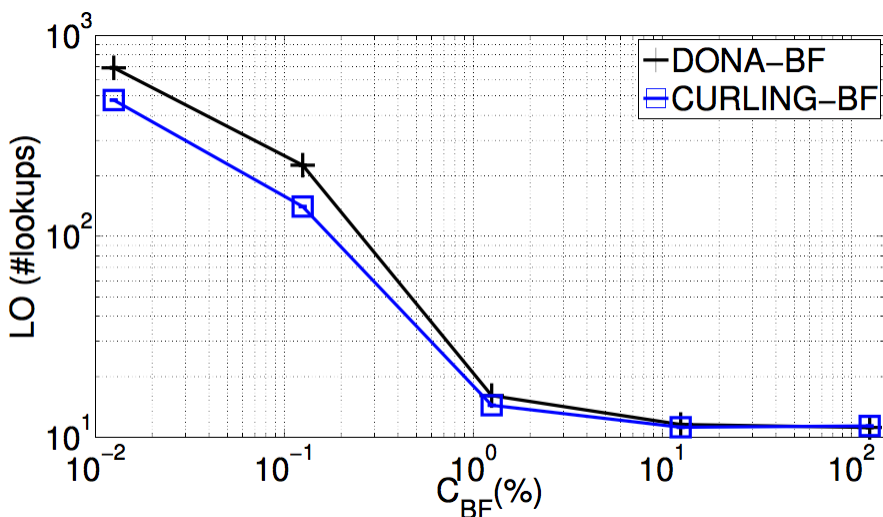
Simulation results (Scaled down topologies)

State size



Simulation results

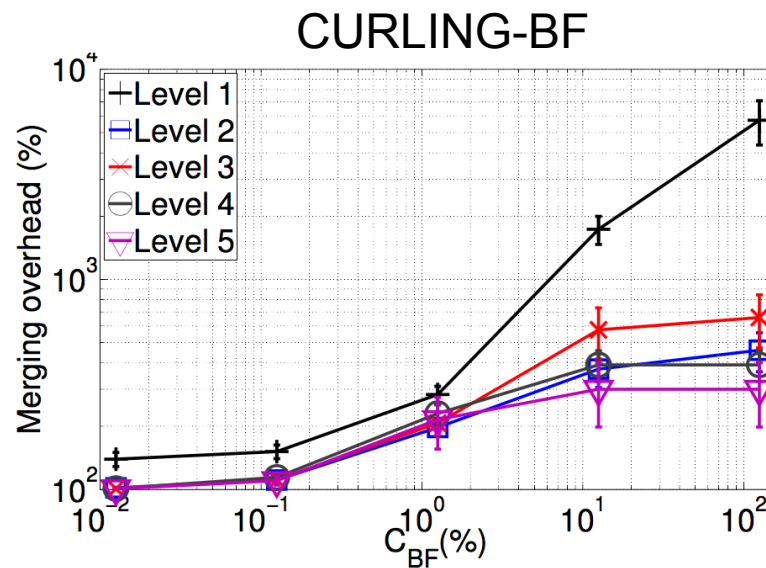
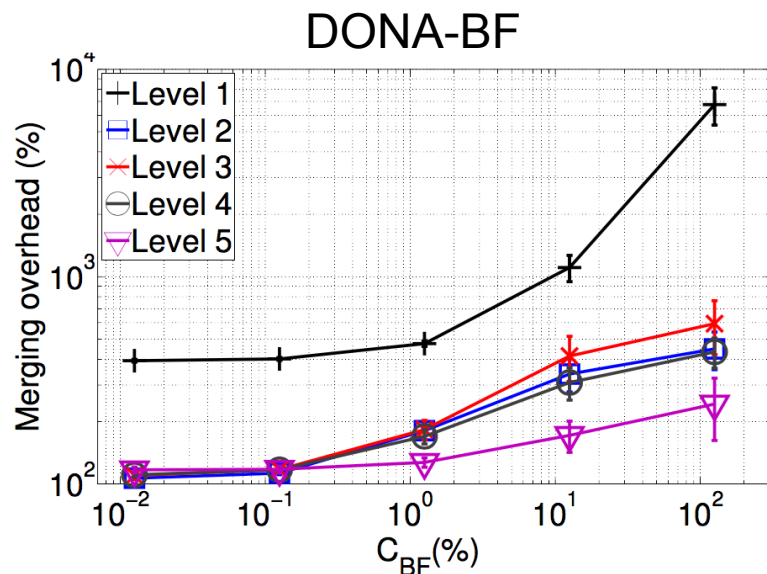
Lookup overheads



- Increased overheads for CURLING
 - Effect of not using peering links
- Low sensitivity to C_{BF}
 - Impact of topology structure (see next)

Simulation results

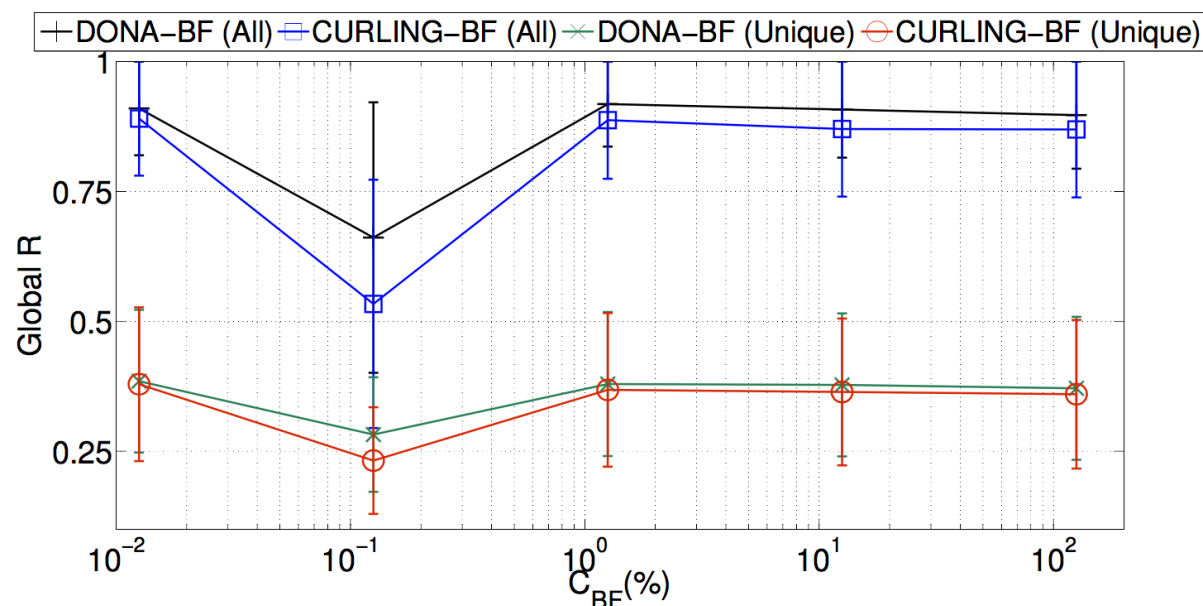
BF merging



- Larger overheads for top tier domains
 - Large number of Stub domains direct customers of top tier domains
 - Non-optimal merging
- Larger overheads for larger C_{BF} values
 - F : lower bound estimation

Simulation results

Global False Positive Ratio

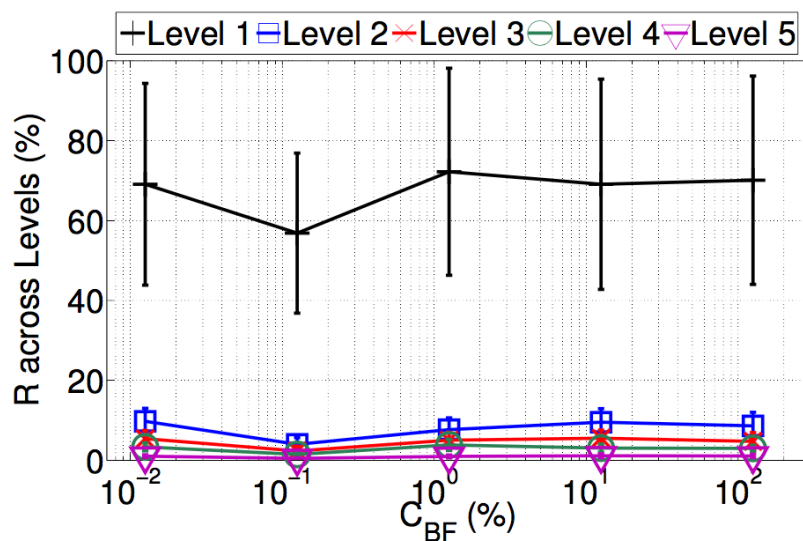


- Extremely high False Positive Ratio
 - Zipf-like workload i.e., multiple requests for popular items
 - Considerably lower ratio for *unique* requests, but still unacceptable
 - No BF update mechanism...

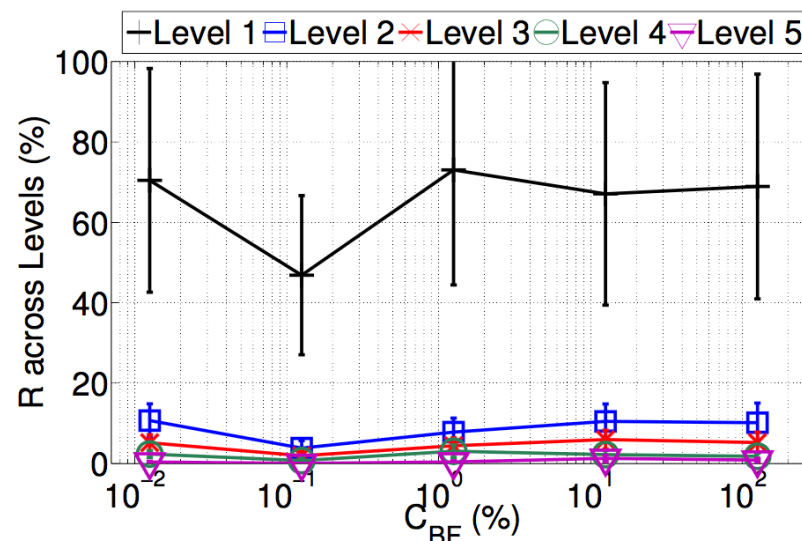
Simulation results

Global False Positive Ratio

DONA-BF



CURLING-BF



- Vast majority of False Positives at tier-1 domains
 - Large concentration of Stub domains at level 2
 - BFs maintained per customer, not merged

Conclusions

- No single BF configuration can lower both memory and processing requirements for all ASes
 - Reducing memory resource requirements for the majority of ASes inflates processing requirements at Tier-1
- Direct connectivity of large content provider ASes to tier-1 inflates False Positives

Future Work

- Uniform Recursive Tree (UTR) model
- Scalable BFs, Dynamic BFs, (d-left) Counting BFs, Cuckoo filters

Thank you.
Questions?



Konstantinos V. Katsaros
k.katsaros@ucl.ac.uk

BACKUP SLIDES

Evaluation Framework

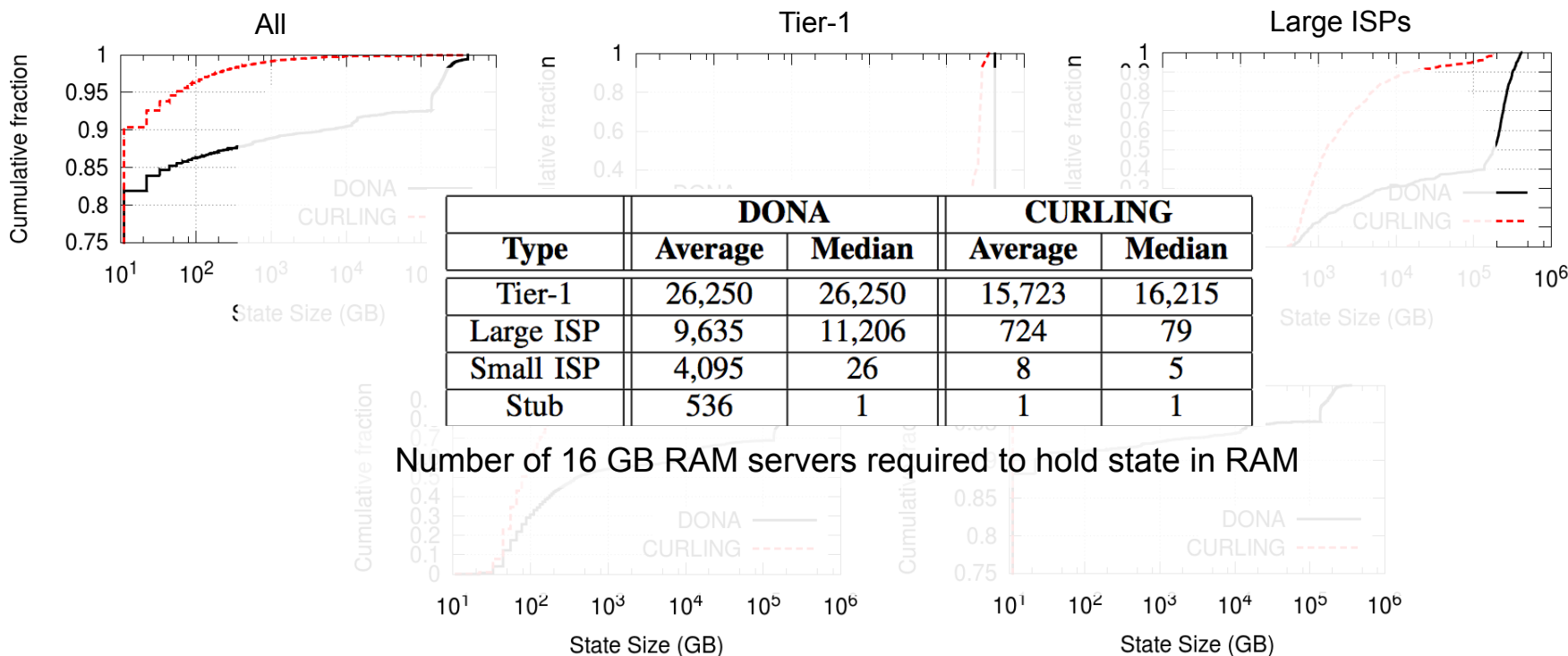
- **Metrics**
 - State size
 - Processing overheads
 - Resolution delay
 - Registration overhead
- **Workload**
 - IO names uniformly distributed across Content Providers at leaf ASes
 - Full-scale: 10^{13} IOs
 - Scaled down: $\sim 10^6$ IOs
- **Network topology**
 - Full-scale: CAIDA trace set
 - $> 45K$ ASes, $> 150K$ links
 - Cone size (c): number of downstream customers
 - Tiers
 - Tier-1
 - Large ISPs ($50 < c$)
 - Small ISPs ($5 < c < 50$)
 - Stub networks ($c \leq 5$)
 - Scaled down topologies

Traffic type	Data plane fraction	Control plane fraction	Object size median	Object size distribution	Object popularity distribution
Web	35.10%	36.40%	10.386 KB	Lognormal-Pareto	Zipf
P2P	15.85%	$2.56 \times 10^{-4} \%$	650.11 MB	Sampling	Mandelbrot-Zipf
Video	19.54%	$37.04 \times 10^{-3} \%$	7.6 MB	Concatenated Normal	Weibull
Other	29.51%	63.57%	5 KB	Normal	Zipf

GlobeTraff traffic mix

Results

State distribution, full-scale, per tier, DONA vs. CURLING

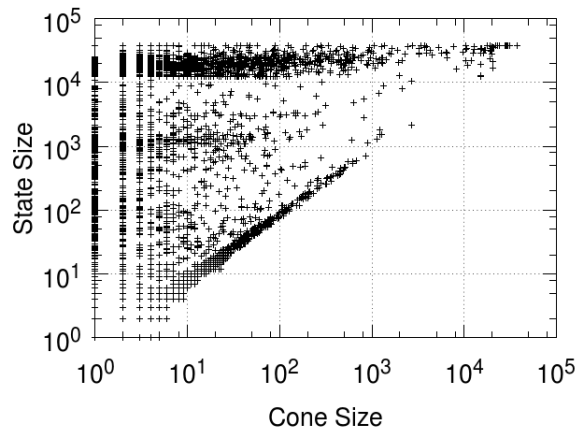


- 420 TB of state for Tier-1 AEs in DONA, for 10^{13} IOs
- < 10 TB for 90% of Large ISPs (CURLING), 30% (DONA)
- Substantially less state for CURLING ...

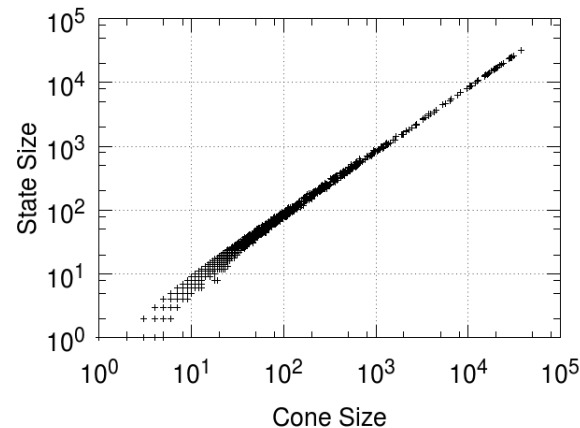
DONA/CURLING Scalability

Large impact of peering links

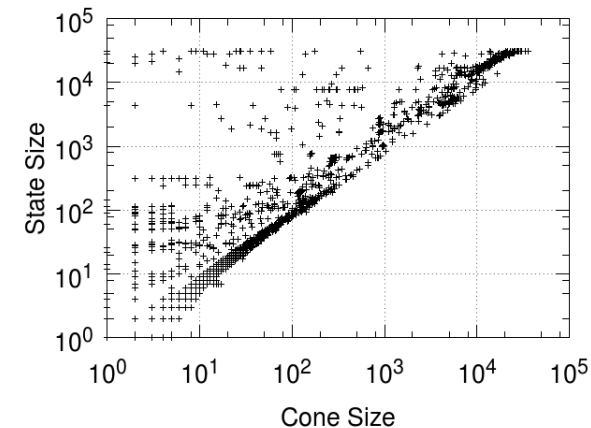
DONA, CAIDA 2013 traces



CURLING, CAIDA 2013 traces



DONA, CAIDA 2011 traces

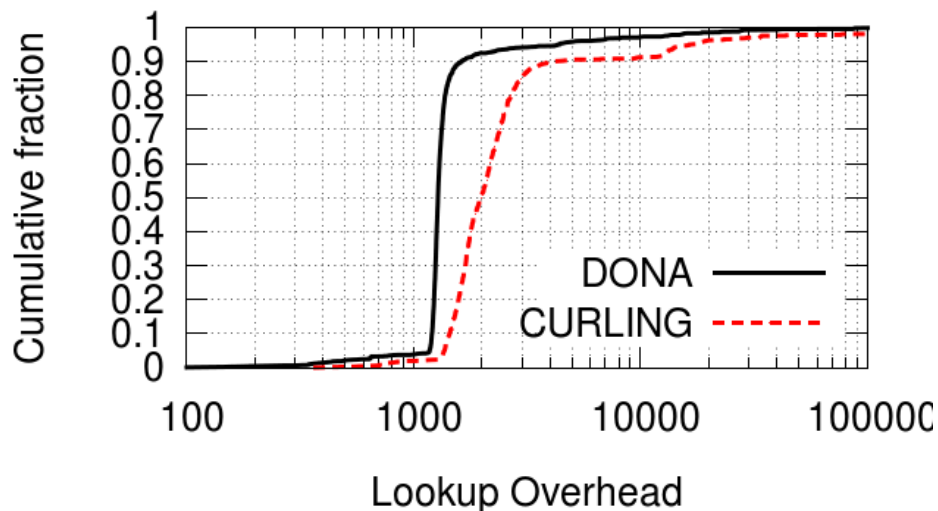


- Not exchanging state across **peering links** (CURLING):
 - Reduces state overheads: 62-fold on average, up to 679-fold for Stub domains
 - Reduces registration traffic: 684% on average
 - Increases processing overhead (2.78-fold on average), especially at top-most domains
 - Increases name resolution paths: 16% on average

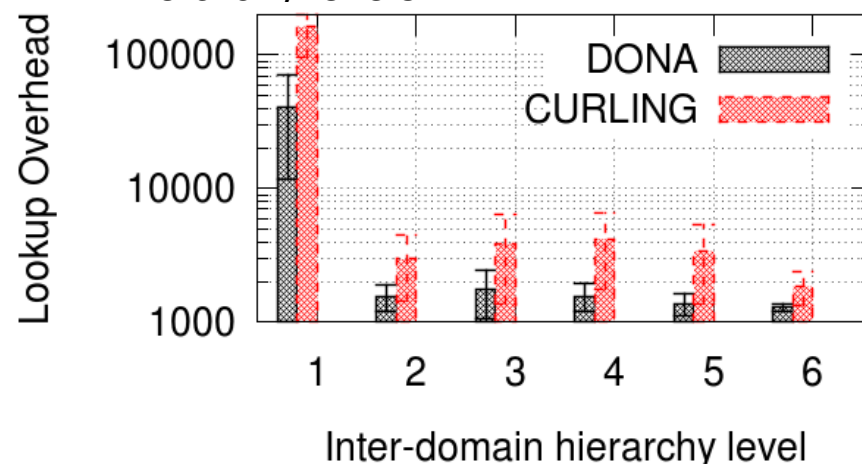
Results

Processing overheads, scaled down, DONA vs. CURLING

Cumulative lookup overhead



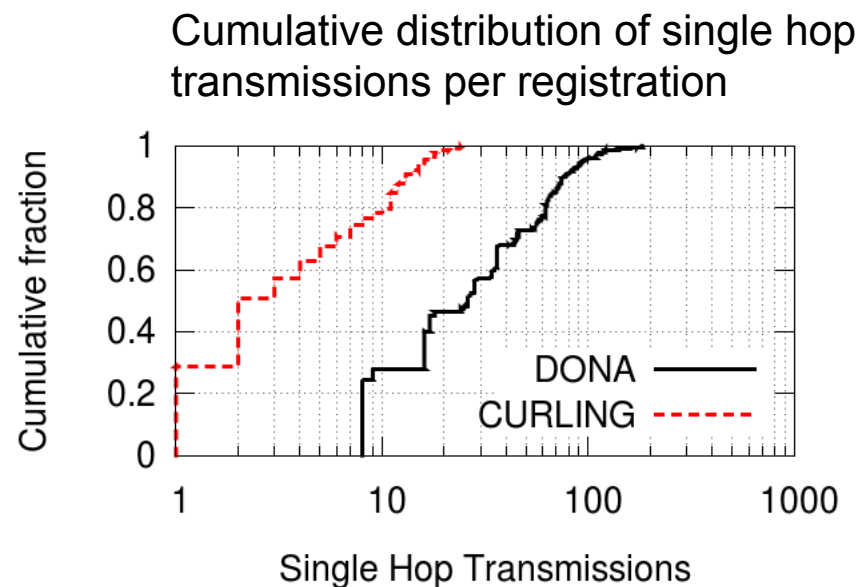
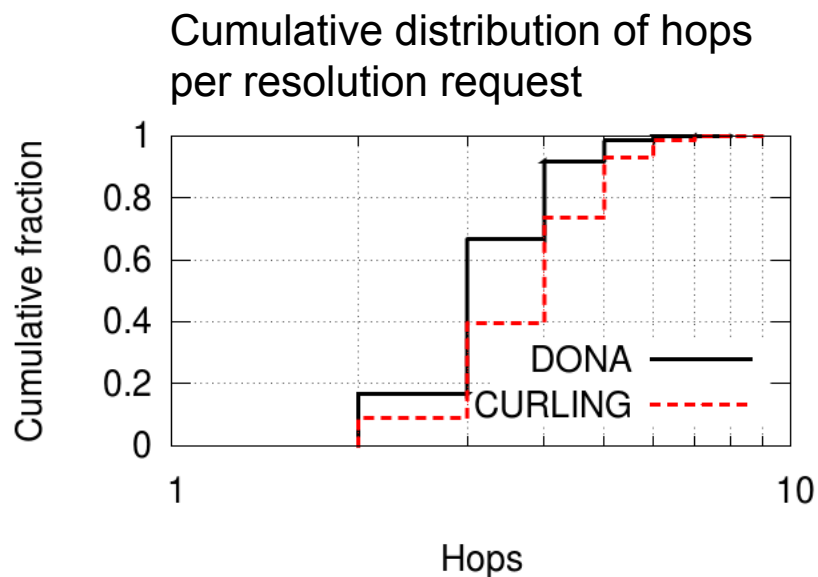
Distribution of lookup overhead across hierarchy levels



- Increased overhead for CURLING (2.78-fold on average)
 - Especially for top-most ASes (3.9-fold on average)
- Not searching for contents in peering domains propagates requests further up

Results

Resolution delays, Registration overheads, DONA vs. CURLING



- Shorter name resolution paths for DONA (16% on average)
 - More resolution requests reach the higher tiers
- 684% increase of registration traffic for DONA
 - Impact of peering links

Further/detailed results

Type	DONA		CURLING	
	Average	Median	Average	Median
Tier-1	26,250	26,250	15,723	16,215
Large ISP	9,635	11,206	724	79
Small ISP	4,095	26	8	5
Stub	536	1	1	1

Number of 16 GB RAM servers required to hold state in RAM

Type	DONA		CURLING		
	Average	Median	Average	Median	Avg. gain
All Tiers	3.778%	0.003%	0.060%	0.003%	62.97%
Tier-1	100.00%	100.00%	59.895%	61.769%	1.67%
Large ISP	36.701%	42.687%	2.758%	0.298%	13.31%
Small ISP	15.599%	0.097%	0.029%	0.018%	537.90%
Stub	2.039%	0.003%	0.003%	0.003%	679.67%

State size per AS expressed as a percentage of the total state size throughout the inter-network (%)

Further/detailed results (Cont.)

Registration traffic (average/median)

	DONA		CURLING	
	Average	Median	Average	Median
Tier-1	66.13 Gbps	66.13 Gbps	39.61 Gbps	40.85 Gbps
Large ISPs	24.28 Gbps	28.23 Gbps	1.82 Gbps	197.08 Mbps
Small ISPs	10.32 Gbps	/ 64.15 Mbps	19.18 Mbps	11.90 Mbps
Stub networks	1.35 Gbps	1.98 Mbps	1.98 Mbps	1.98 Mbps

10^{13} IOs, two weeks average IO lifetime, REGISTRATION size = 1 KB.