UNIVERSITY OF CAMBRIDGE

TecVis

Information-centric Internetworking

A Few Insights across the Board

Thanks to George Parisis, Ben Tagger, Stuart Porter, Jimmy Kjällman, Martin Reed and Mays Al-Naday for their contributions!

Dirk Trossen, Co-Founder TecVis LP
& Visiting Researcher (Computer Laboratory)

Overview

- Background and motivation
- Architectural foundations
  - Few more details on:
    - intra-domain forwarding
    - Inter-domain forwarding
    - Network coding
- Application: personalised media delivery
- Prototype and test bed

We All Know About Video: Staggering Numbers

- Over 4 billion hrs of videos watched on YouTube every month
  - 72 hrs uploaded on YouTube every minute
  - 70% of traffic from outside US
- The 2012 Olympics broke all records
  - BBC delivered 2.8 petabytes on its busiest day, 700Gb/s during the B. Wiggins' gold
- 74 mins average BBC iPlayer TV usage per week
  - 1.6 mio daily iPlayer viewers in July 2011
- …in all this, mobile usage just started to take off!
  - YouTube mobile traffic tripled in 2011

…With Staggering Forecasts (Cisco)

- Annual global IP traffic will reach the zettabyte threshold by 2015
- The average smartphone will generate 1.3 GB of traffic per month in 2015 (26x)
- In 2015, there will be 6 million Internet households worldwide generating over a terabyte per month in traffic
- By 2012 Internet video will account for over 50 percent of consumer Internet traffic

…But There is so Much More Than Content – It's Information!

Mobile Phones

Sensors

Things

Your Personal Photos

Health Data

Your Personal Life

Retail Data

The Internet Has Always Been About Information – And It Copes Well With It!

That is correct… (to a point to be discussed)

BUT: Economics have changed the possible starting points for a design

- Computing and storage resources are NOT scarce anymore
  - This led to an almost ubiquitous availability of processing and memory
- Information availability has changed attitude of users
  - WHAT is primary, WHO and WHERE mostly secondary!
  - Information is often not locked anymore behind portals

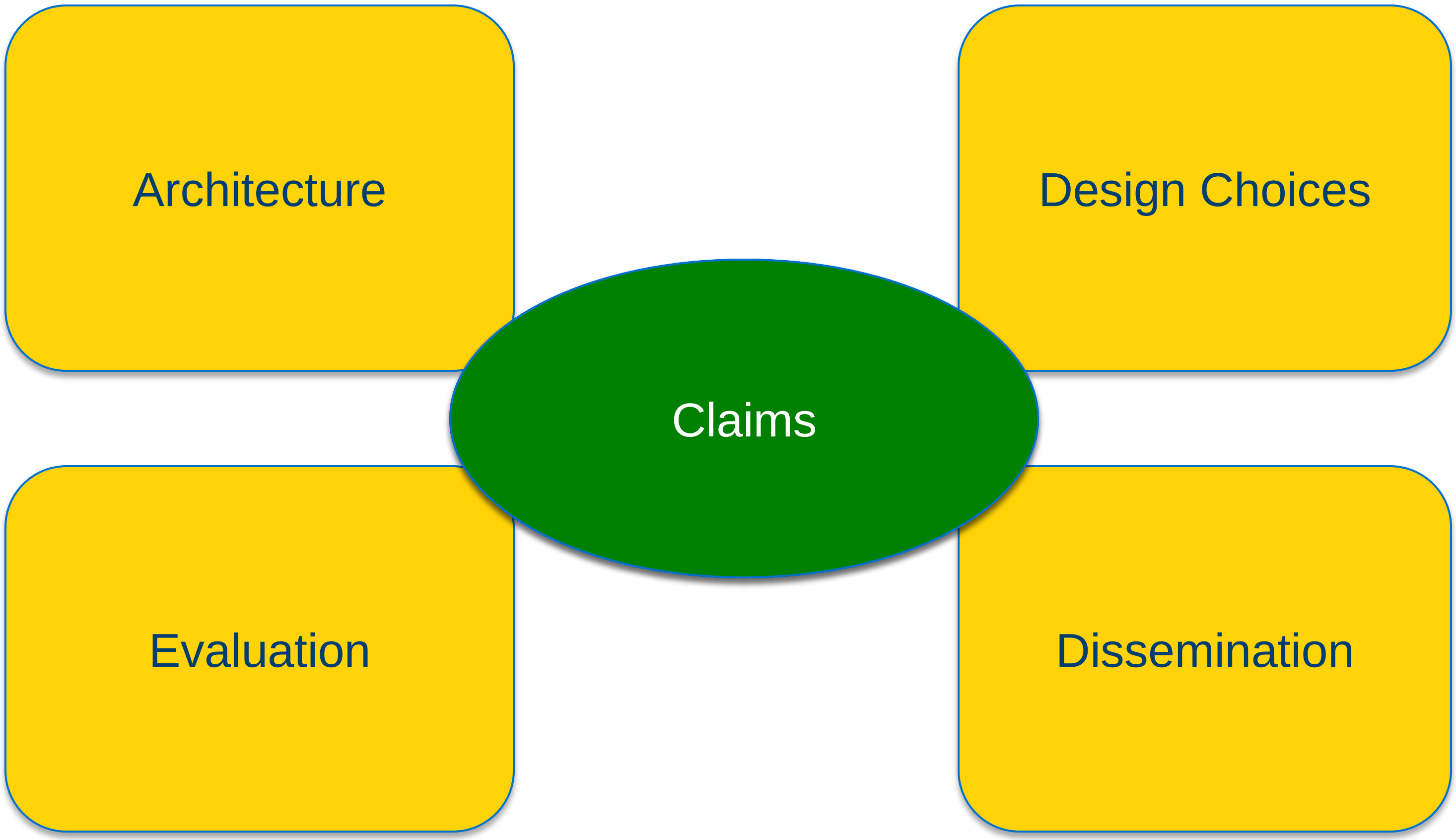
⇒There is desire to fully optimize the usage of resources (wherever they are located)

Hypothesis

*A systems approach that operates on **graphs** of **information** with a **late** (as late as possible) binding to a location at which the **computation** over this graph is going to happen, enables the full potential for **optimization**!*

This systems approach requires to marry information & computation (and with it storage) into a single design approach for any resulting distributed system

Main Challenges



Our Claims: As Formulated So Far*

Design, develop and evaluate a novel information-centric pub/sub-based internetworking architecture that

- Provide an improved *impedance match* towards application-level concepts
- Provide *tussle delineation* of crucial functions
  - Tussles here refer to conflicts between stakeholders
- Enable *optimization* of sub-architectures
- Provide high *performance*
- *Scale* to the needs of the Future Internet

* see CCR 04/2010: Arguments for a new Information-centric Internetworking Architecture

Our Main Challenges: Architecture

Provide a sound architectural framework for information-centric networking

Main thrusts:

- Design tenets and their specific or general viability
- Translate tenets into coherent set of concepts
  - Provide a set of coherent architectural arguments for their viability
  - In particular the proper (socio-)economic arguments

Our Main Challenges: Design Choices

Develop a set of design choices to support our architectural claims

Main thrusts:

- Rendezvous throughout all (recursive) levels of the architecture
- Inter-domain topology formation
- Topology management (focus on optical and wireless)
- Forwarding
- Caching & Transport
- Information-centric middleware solutions

Our Main Challenges: Evaluation

Provide the required proofs for our architectural claims

Main thrusts:

- Implementation (prove that it runs – and performs)
- Simulation (prove that it scales – and performs)
- Socio-economics (prove that its design is viable)
- Economics (prove that it is economically sensible)

Our Main Challenges: Dissemination

Provide the required tools for disseminating our results

Main thrusts:

- Implementation (a tool to create a community)
- Test bed (a place to meet and try out)
- Website (a place to exchange)
- Course material (a tool to educate the new generation)
- Exploitation strategies (a tool to convince the stakeholders)

Publications and presentations are means to an end for all the above

Hypothesis

*A systems approach that operates on **graphs** of **information** with a **late** (as late as possible) binding to a location at which the **computation** over this graph is going to happen, enables the full potential for **optimization**!*

How to go about this?

Starting Point: Solving Problems in Distributed Systems

- One wants to solve a problem, each of which might require solving another problem
  - Examples:
    - Send data from A to B(s), involving fragmentation along the link(s)
    - Disseminate a video over a local network
- Problems involve "*a collection of information that*" an implementation *"can use to decide what to do"*, which is to implement a problem solution (*)

-> Computation in distributed systems is all about *information dissemination* (pertaining to a task at hand)

*REF: S. J. Russell, P. Norvig, "Artificial Intelligence: A Modern Approach", 2nd Edition, Pearson Educ., 1998

Desired System Properties…

- **Manipulation of (structured) information flows for computational purposes**
  - Expose service model and provide late binding (*WHAT->WHO*)
- **Modularity within a single computational problem**
  - Provide modular core functions (*enable optimization*)
- **Modularity across computational problems**
  - Provide rigorous but flexible layering (*deconstrain constraints*)

REF: CHIANG, M., LOW, S. H., CALDERBANK, A. R., AND DOYLE, J. C. Layering as Optimization Decomposition: A Mathematical Theory of Network Architectures. Proceedings of the IEEE (2007)

…Translated into Design Tenets…

- Provide means for identifying individual information (items)
  - Can be done via labeling or naming
- Provide means for scoping information
  - Allows for forming DAGs (directed acyclic graphs)
- Expose service model
  - Can be pub/sub
- Expose core functions
  - Rendezvous, topology management, and forwarding
- Common dissemination strategy per sub-structure of information
  - Define particulars of functional implementation and information governance

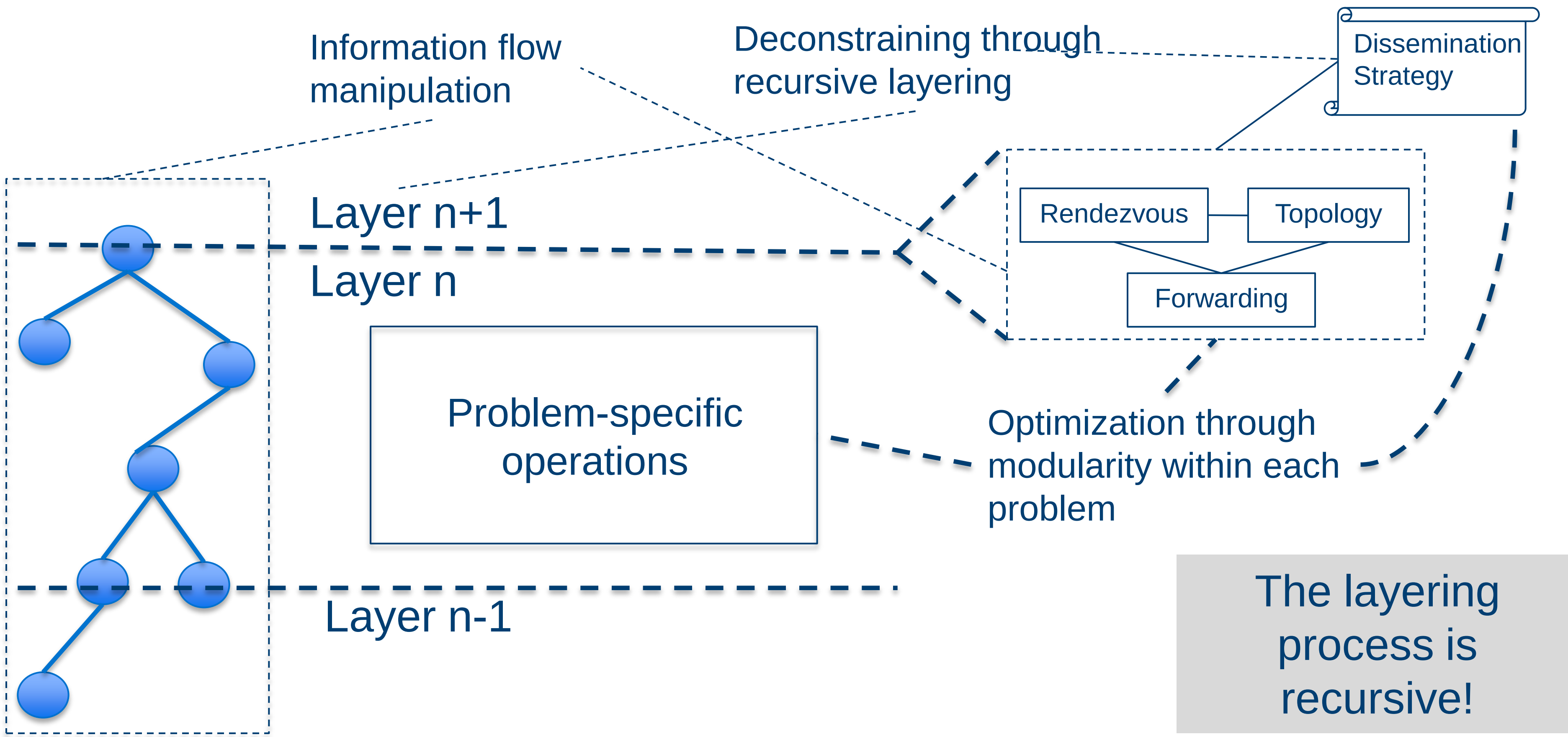
…With An E2E Principle…

The problem in question can be implemented through an assembly of sub-problem solutions, whose individual dissemination strategies are not in conflict with the ones set out by the problem in question.

- Hence, problems are assembled to larger solutions by recursively applying the scoping invariant of the functional model!
- Conflicts are avoided through design and re-design, e.g., via standards procedures!
- Can extend this to runtime reconciliation!

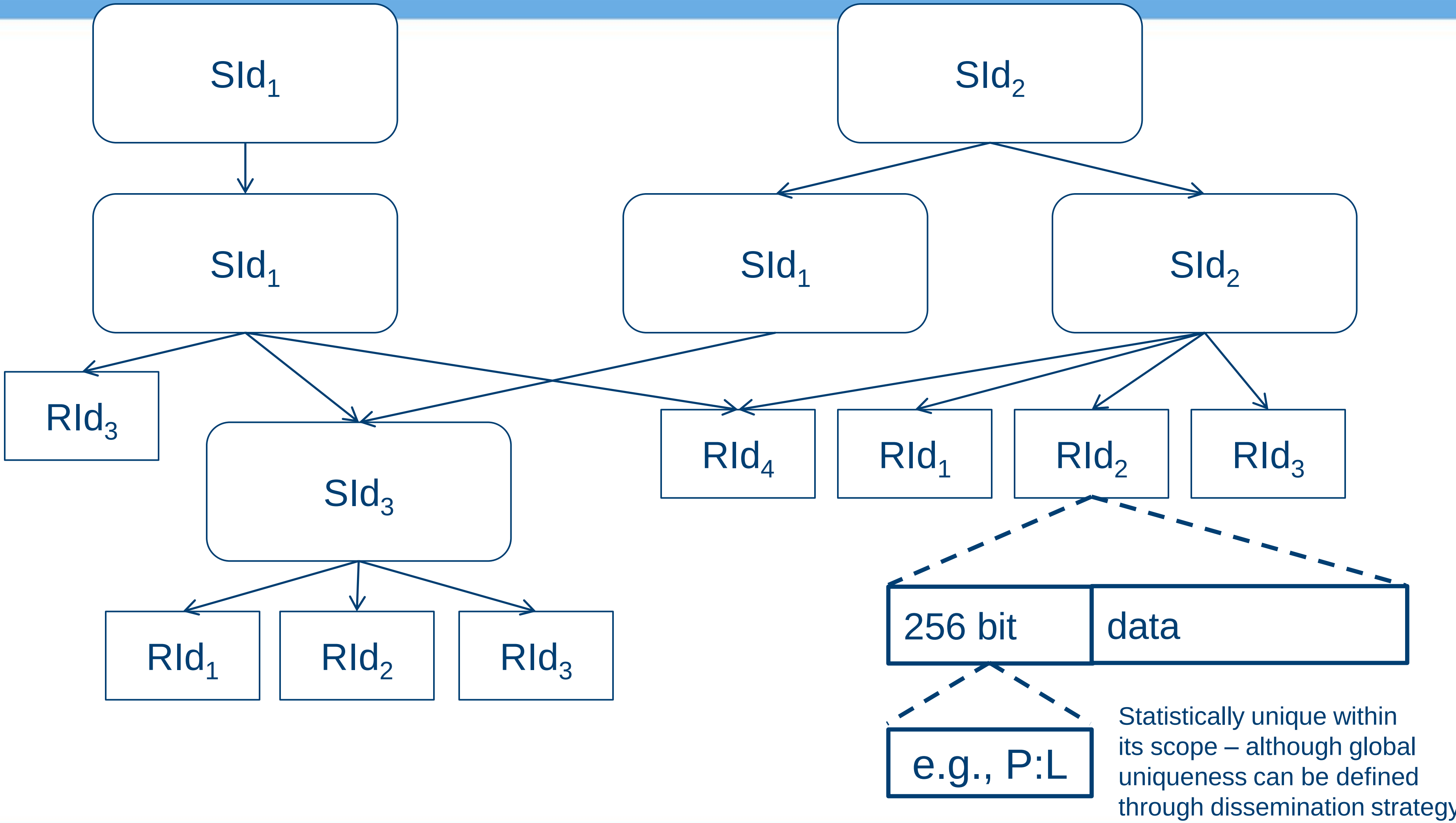
NOTE: I leave it as a thought exercise to relate this to the IP E2E principle!

…And Placed into a Layered Model



REF: DAY, J. Patterns in Network Architecture - A Return to Fundamentals. Prentice Hall, 2008

Operating on Graphs of Information

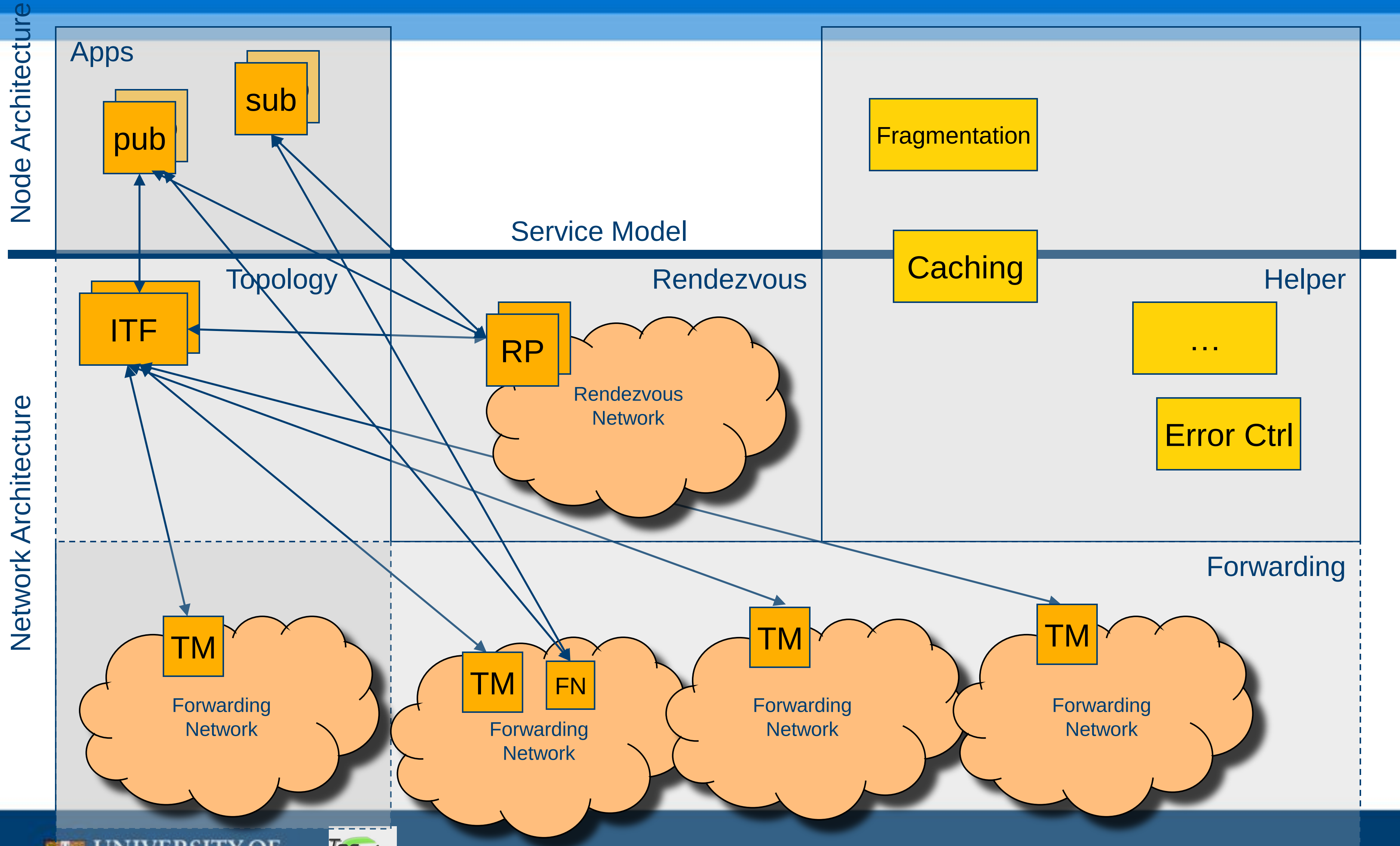


Information Semantics: Immutable vs. Mutable Items

- Documents
  - Each RId points to immutable data (e.g., document version)
  - Not well suited for real-time type of traffic
  - Each item is identifiable throughout the network
- Channel
  - Each RId points to channel of data (e.g., a video stream), i.e., the data is mutable
  - Well-suited for video type of traffic
  - Problems with caching though (since no individual video segments visible)

…Coming Together in A Global Architecture

RP : Rendezvous point
ITF : Inter-domain topology formation
TM : Topology management
FN : Forwarding node



Example of One Core Function

Forwarding with Built-in (Native) Multicast Capability

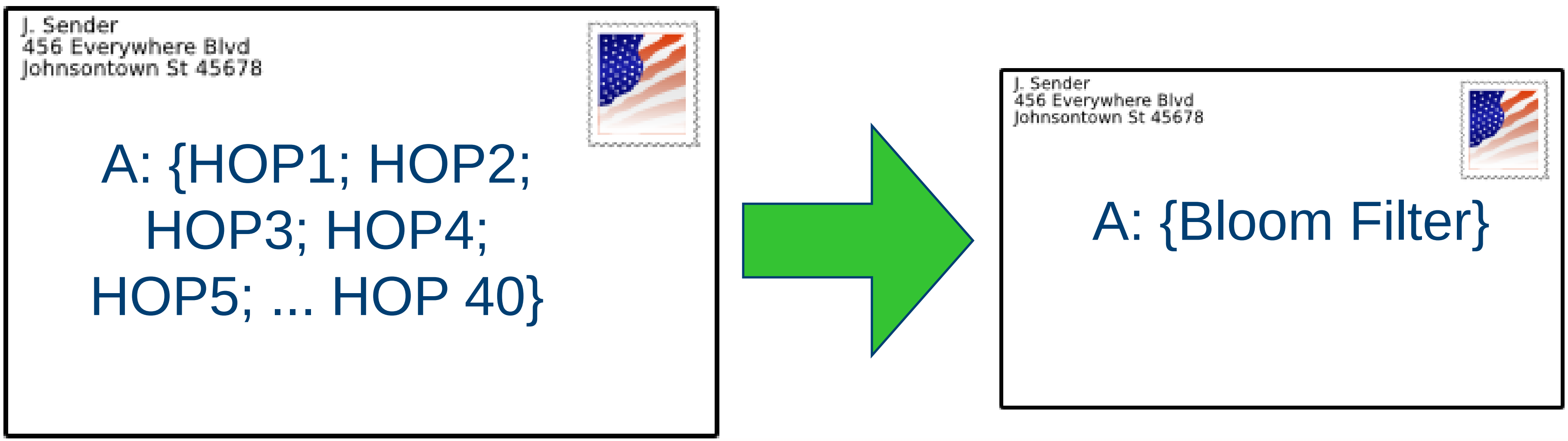
Motivation

Information is sent along a route of (intra-domain) hops in the Internet

-> Requires some form of minimal state in each hop

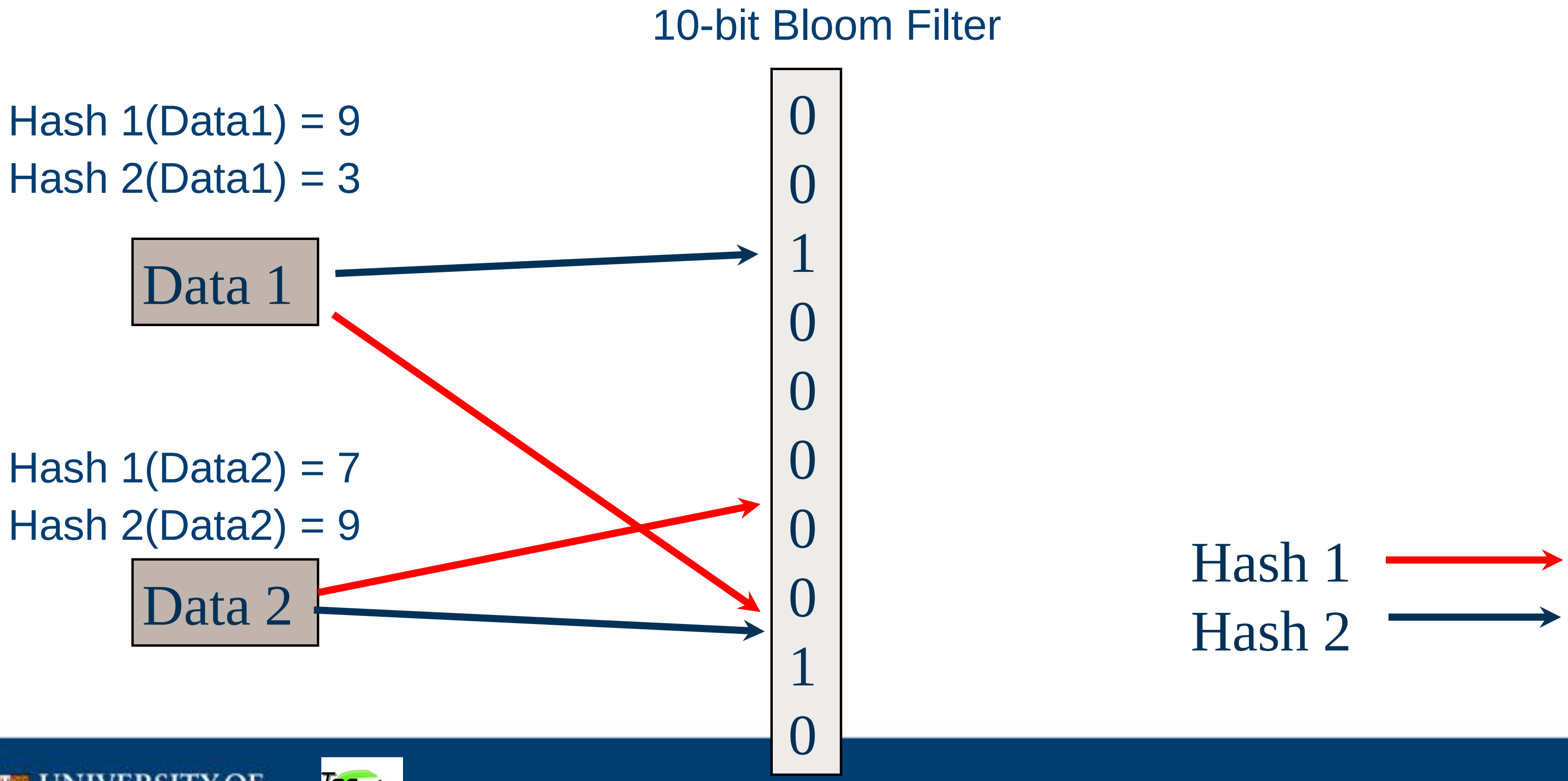
- If forwarding on names, limiting this state is hard/impossible

Question: What if we could instead include the state in the packet?



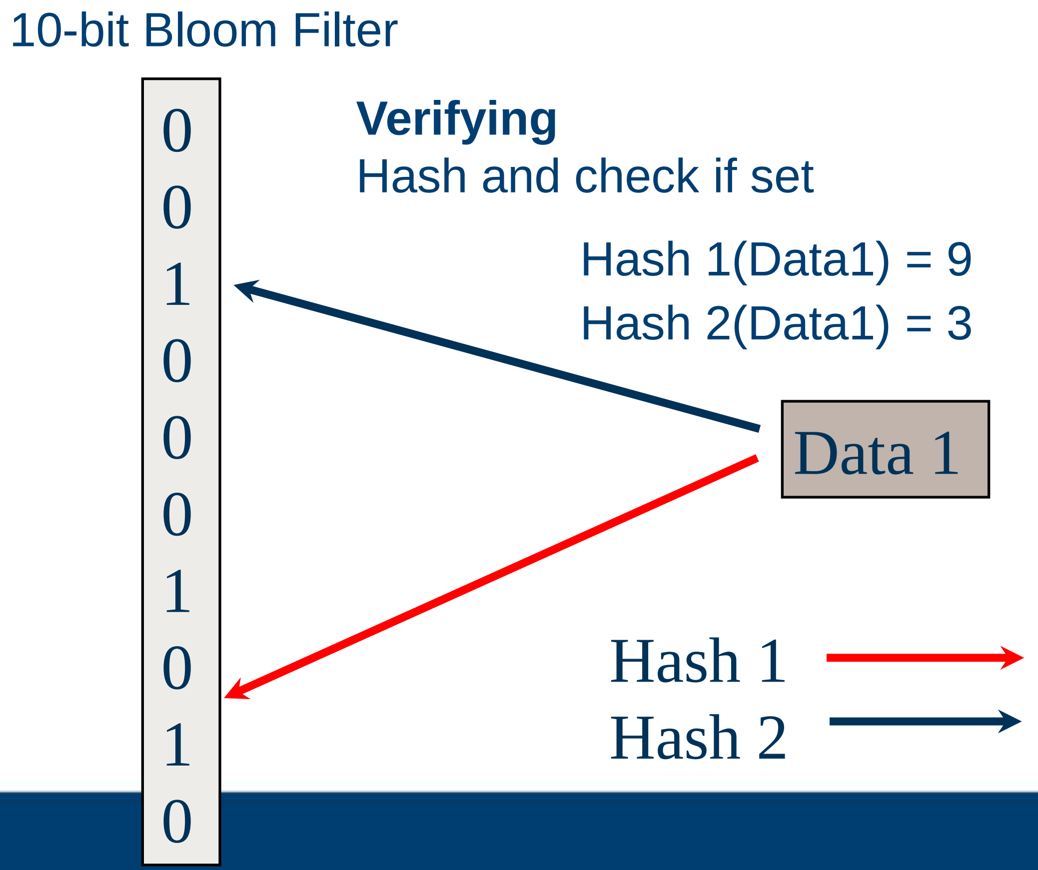
What are Bloom Filters?

- Inserting items
  - Hash the data n times, get index values, and set the bits



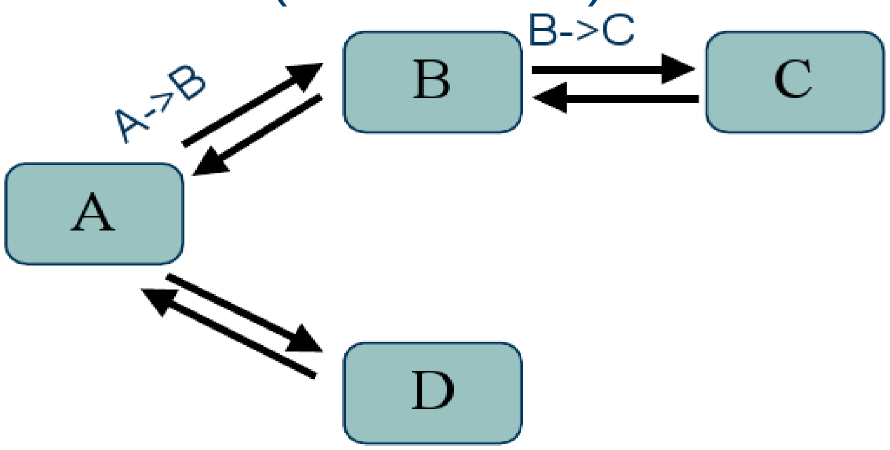
What are Bloom Filters?

- Test if “Data 1” has been inserted in the BF
  - All corresponding bits are set => positive response!



Idea: Line Speed Publish/Subscribe Inter-Network (LIPSIN)

- Line speed forwarding with simplified logic
- Links are (domain-locally) named instead of nodes (LId), therefore there is no equivalent to IP addresses
- Link identifiers are combined in a **bloom filter** (called zFilter) that defines the transit path



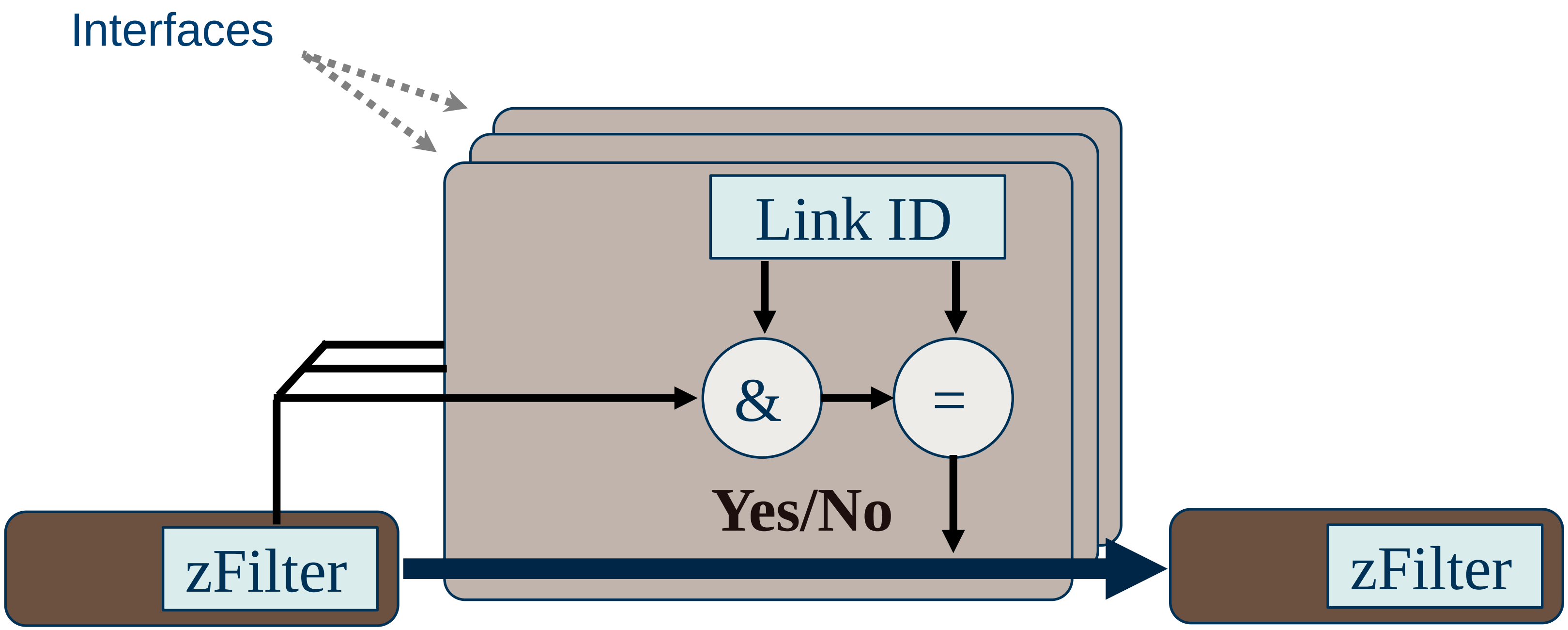
- Advantages
  - Very fast forwarding
  - No need for routing tables
  - Native multicast support

A->B	0	1	0	0	0	1	0	0	1
B->C	1	0	0	0	0	1	1	0	0
zF: A->B->C	1	1	0	0	0	1	1	0	1

UNIVERSITY OF CAMBRIDGE

Forwarding Decision

- Forwarding decision based on binary AND and CMP
  - zFilter in the packet matched with all outgoing Link IDs
  - Multicasting: zFilter contains more than one outgoing links



Problem: False Positives in Forwarding

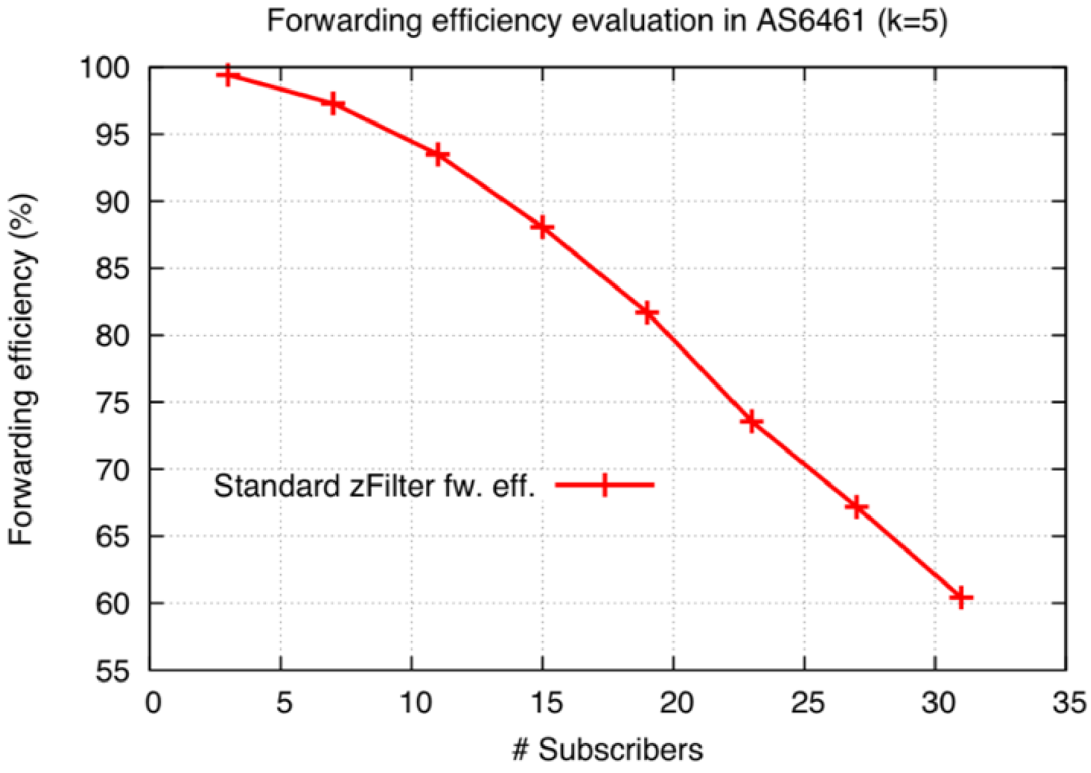
False positives occur when test is positive in a given node despite non-hashed LId (probability for consecutive false positives is multiplicative!)

- Increase with number of links in a domain (since more data is hashed into constant length Bloom filter)
- Two immediate solutions:
  - Use Link Identity Tags: tag a single link with N names instead of one, then pick resulting Bloom filter with lowest false positive probability
  - Virtual trees: fold "popular" sub-trees into single virtual link, i.e., decrease number of LIds to be used

Forwarding Efficiency

- Simulations with
  - Rocketfuel
  - SNDlib
- Forwarding efficiency with 20 subscribers
  - ~80%

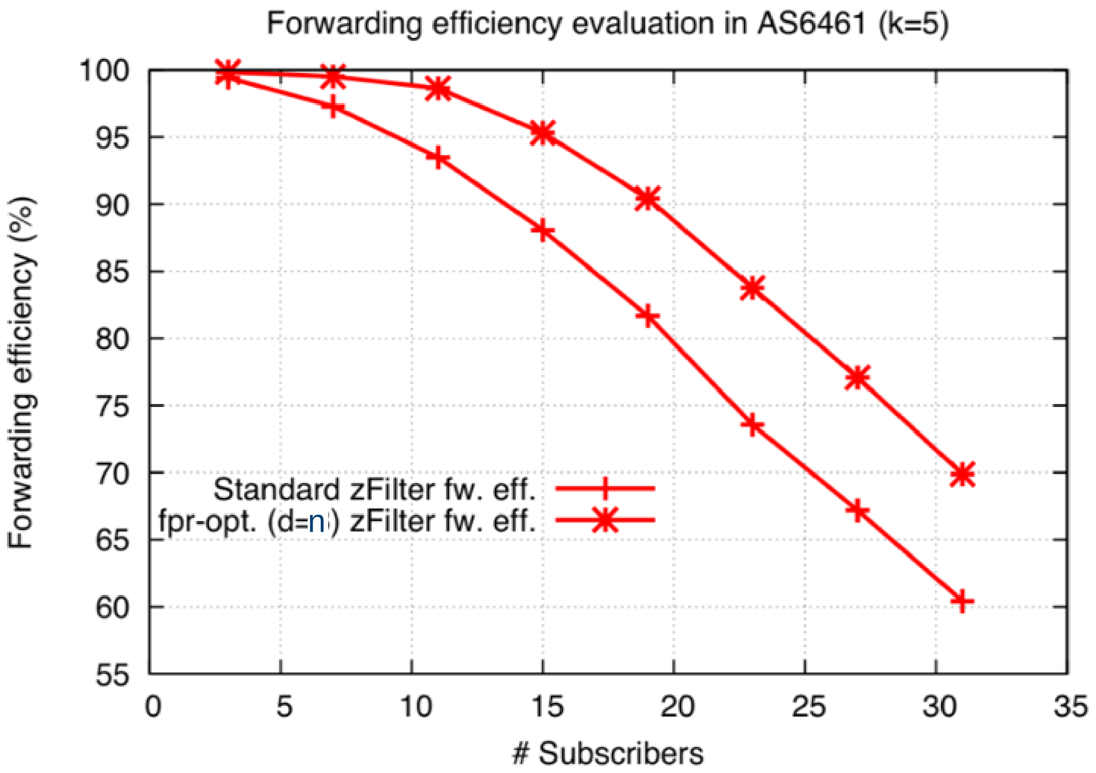
  -> suited for MAN-size multicast groups



UNIVERSITY OF CAMBRIDGE

Forwarding Efficiency

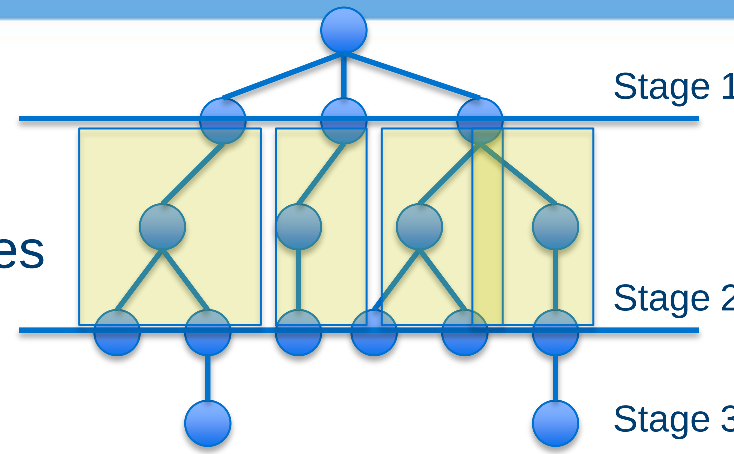
- Simulations with
  - Rocketfuel
  - SNDlib
- Forwarding efficiency with 20 subscribers
  - ~80%
  - Can be optimized to 88% using extended mechanisms



From Efficient Forwarding to Scale

Going Beyond LIPSIN – scaling to any size tree!

Idea: Multi-stage BF Forwarding



- Divide a delivery tree into stages
  - Generally, each stage has individual trees
  - Operation performed at topology manager
- Provide single BF forwarding identifier per stage

<256 bit	data

- Concatenate all stage into variable size header

<256 bit	<256 bit	<256 bit	data

- Perform BF-based forwarding at each stage
- Remove appropriate BF after each stage

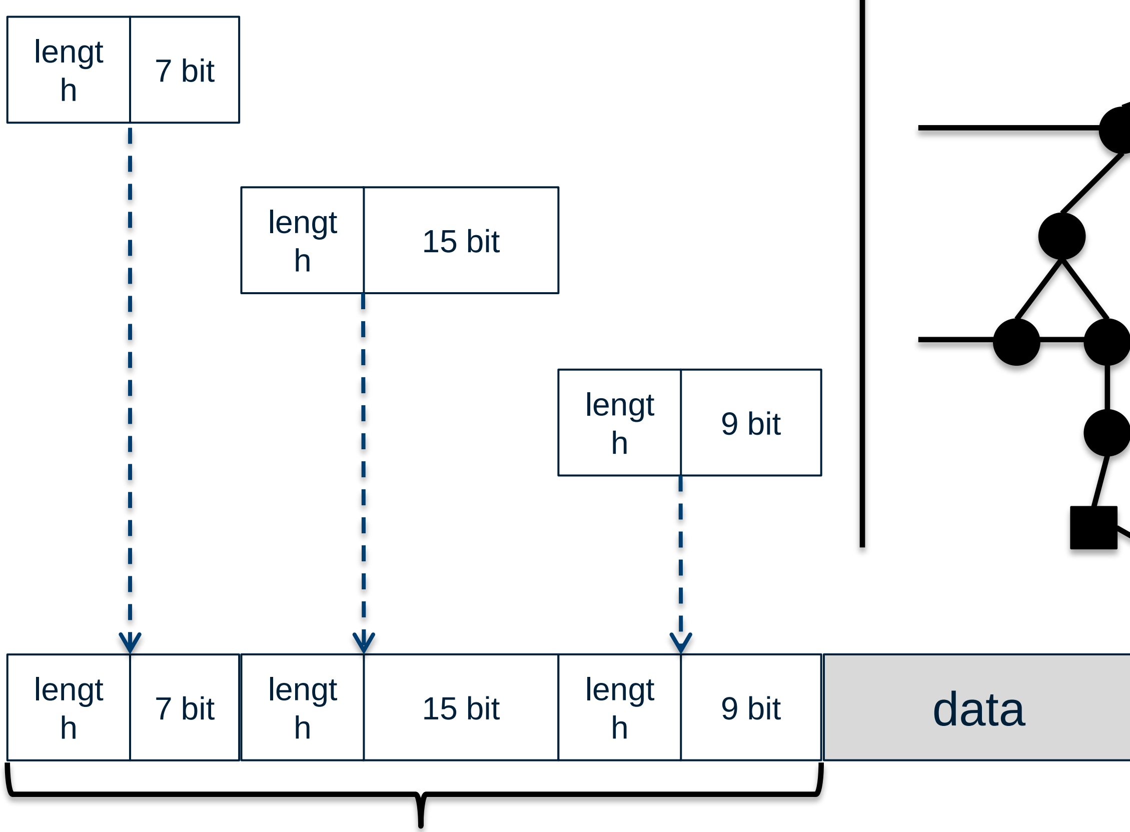
More specifically: Topology Formation

- Calculate tree for given <pub,subs> relation
- For each stage:
  - Define **in_tree** as the set of LIds being in the tree and **out_tree** as the ones not
  - Determine minimal length of BF that can hold **in_tree** with P(false positive)=0 (with the help of **out_tree**)
  - Determine BF through ORing **in_tree** into BF
    - Test if BF would cause false positive (increase, if so)
- Determine overall header through
  - Write length of stageBF through *Elias omega* encoding
  - Write stageBF

For all stages

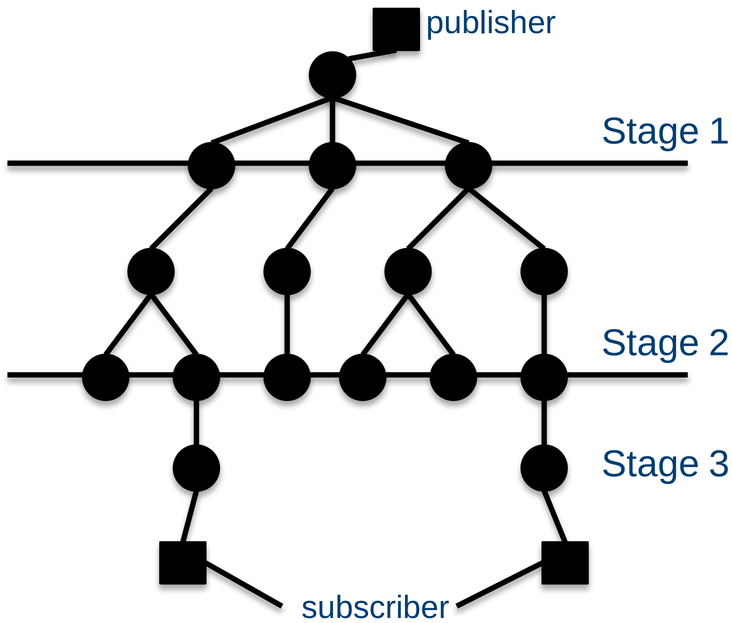
In a Nutshell

Stage-level forwarding identifiers



Final forwarding identifier

Delivery Tree

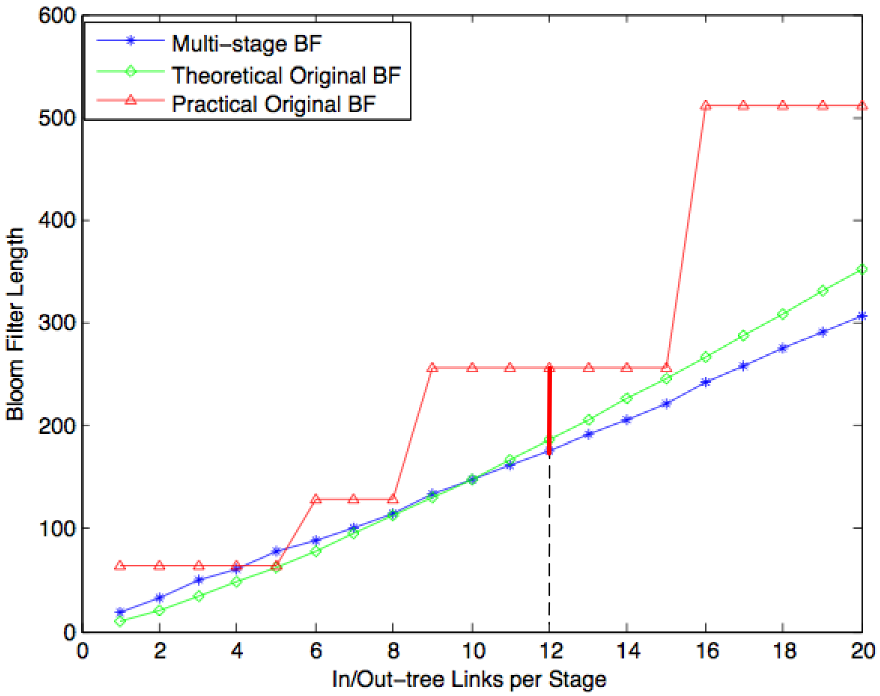


Pros and Cons

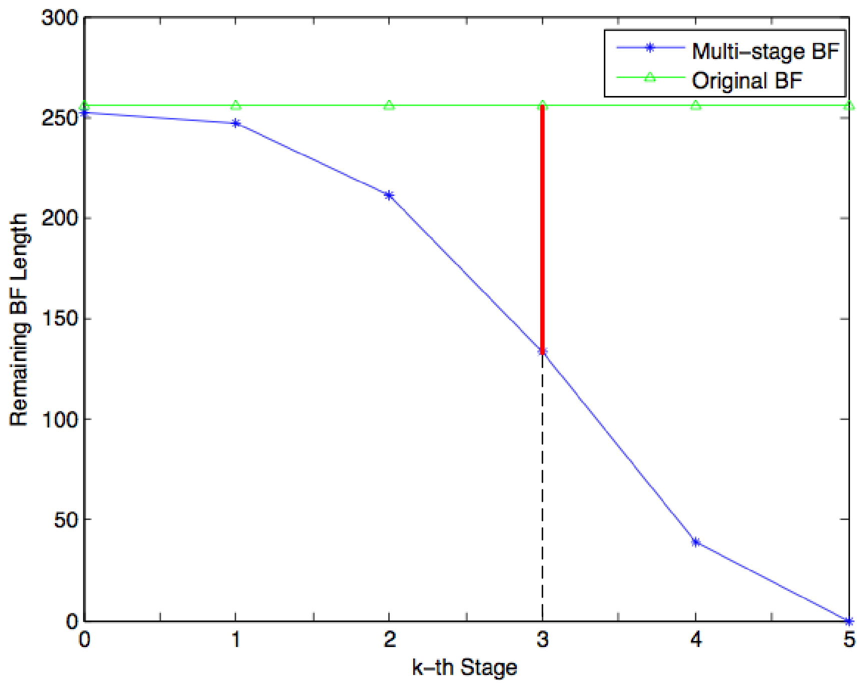
- Advantages
  - Arbitrary tree size (limit only when restricting maximum size for variable length header)
  - **Remove false positive (and all its drawbacks)!**
    - Tradeoff between false positive and header length possible
  - Tradeoff between false positive rate and header size
  - Single hop vs multi-hop stages possible (single hops naturally limit BF anomalies)
  - **Lends itself to inter-domain as well as intra-domain forwarding**
- Disadvantages
  - Higher complexity in forwarding (but only at the stage boundaries)
  - Possibly higher overhead due to variable length, but overhead reduces as you traverse the tree

Header Length: MS-BF vs. LIPSIN

MS-BF vs. LIPSIN (realistic deployments)



Shrinking Header when Traversing Network



Optimising Processing

- BF-based forwarding requires the BF-encoded link identifiers to exist
- Unknown a-priori length of BF-based forwarding identifier requires BF-encoding (i.e., hashing) at runtime

Solution:

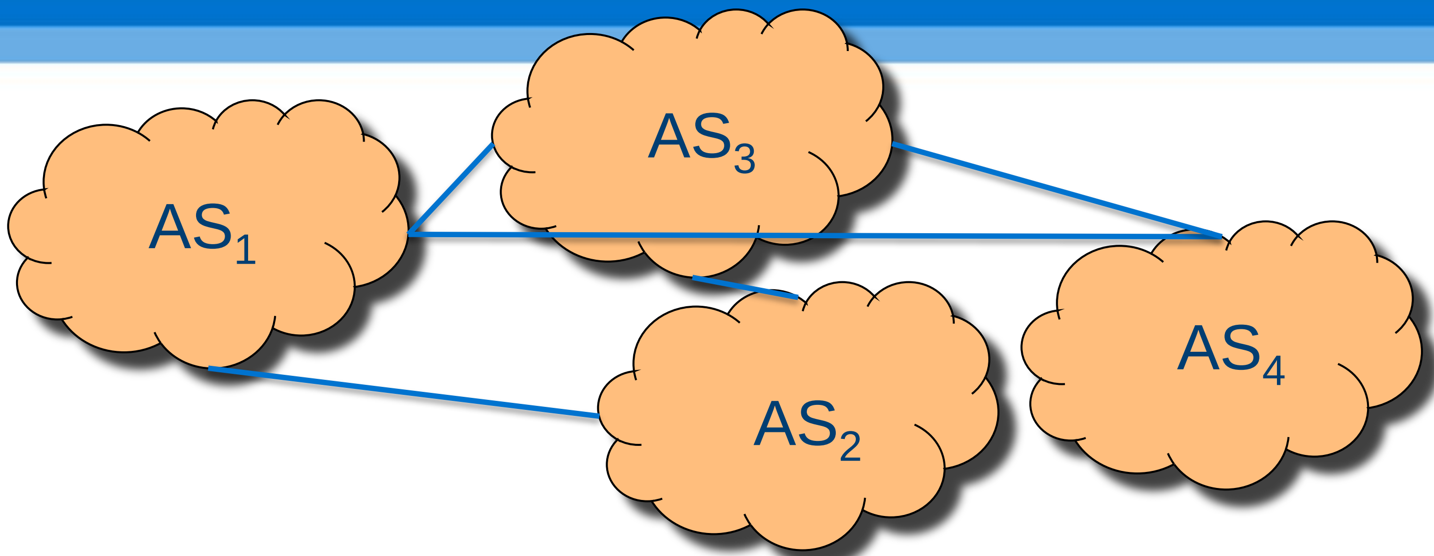
- Use runtime-optimised hashing solutions
- Use **pre-computed hashes** with one hash per BF identifier length
    - Size of locally stored hash table depends on overall connectivity (i.e., length of maximum BF-encoded forwarding identifier)

  -> come to this problem later again!

Reaching the End of the (ICN) World

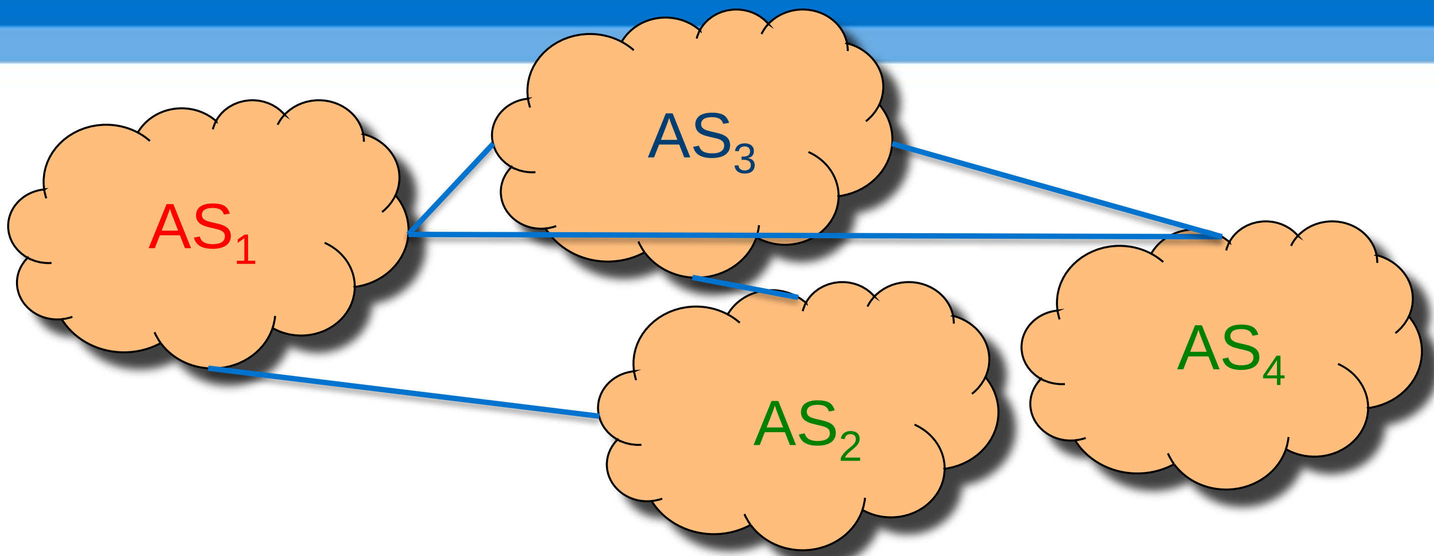
Inter-domain Forwarding in PURSUIT

Assumptions



- Internet comprised of autonomously managed networks (AS)
  - AS-internal resource management, including choice of forwarding is left to individual AS (and essentially arbitrary from the perspective of inter-domain)
- Connectivity between ASes governed by policy contracts
- Partial exposure of these contracts across the ASes

ICN Starting Point



- Information published in AS_1 and subscribed to in AS_2 and AS_4
- Matching of demand (in AS_2 and AS_4) to supply (through AS_1) already done through global rendezvous solution

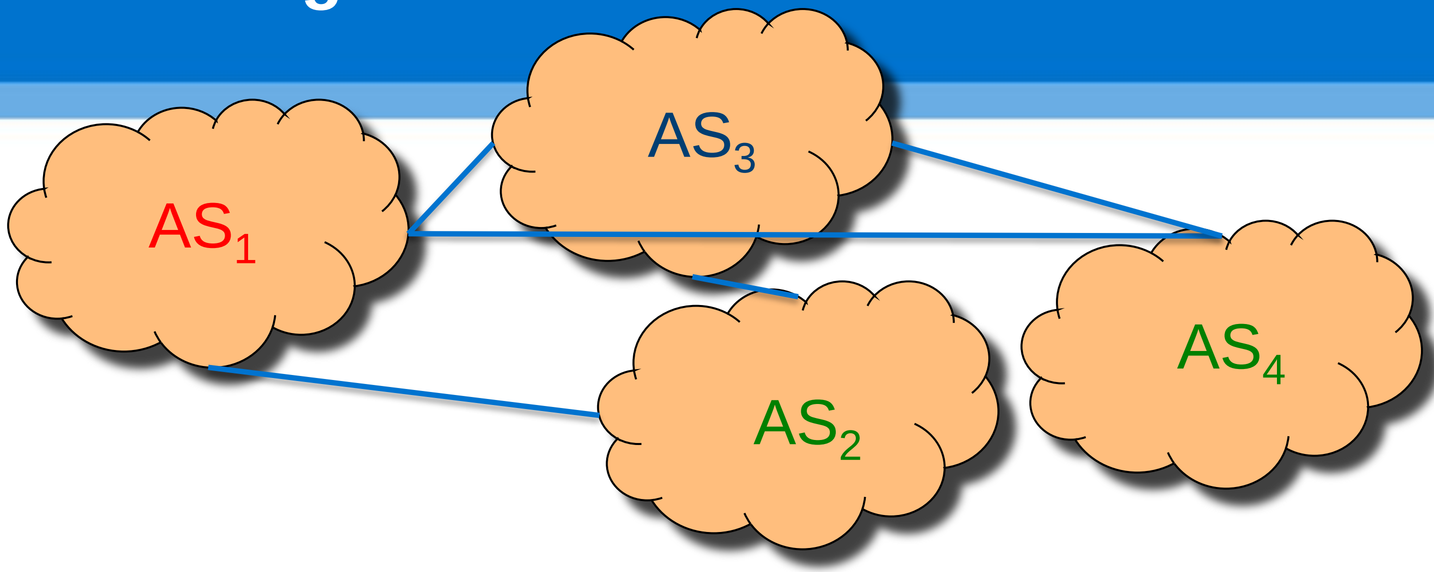
Problem Statement

How do we get the information transferred from AS_1 to AS_2 and AS_4?

Constraints:

1. Provide policy-compliant routing along the agreed upon contracts
2. Support the inherent multipoint notion of pub/sub
3. Do not require network-wide knowledge of AS internals
4. Scale to current Internet connectivity and beyond

Initial Thought: E2E LIPSIN

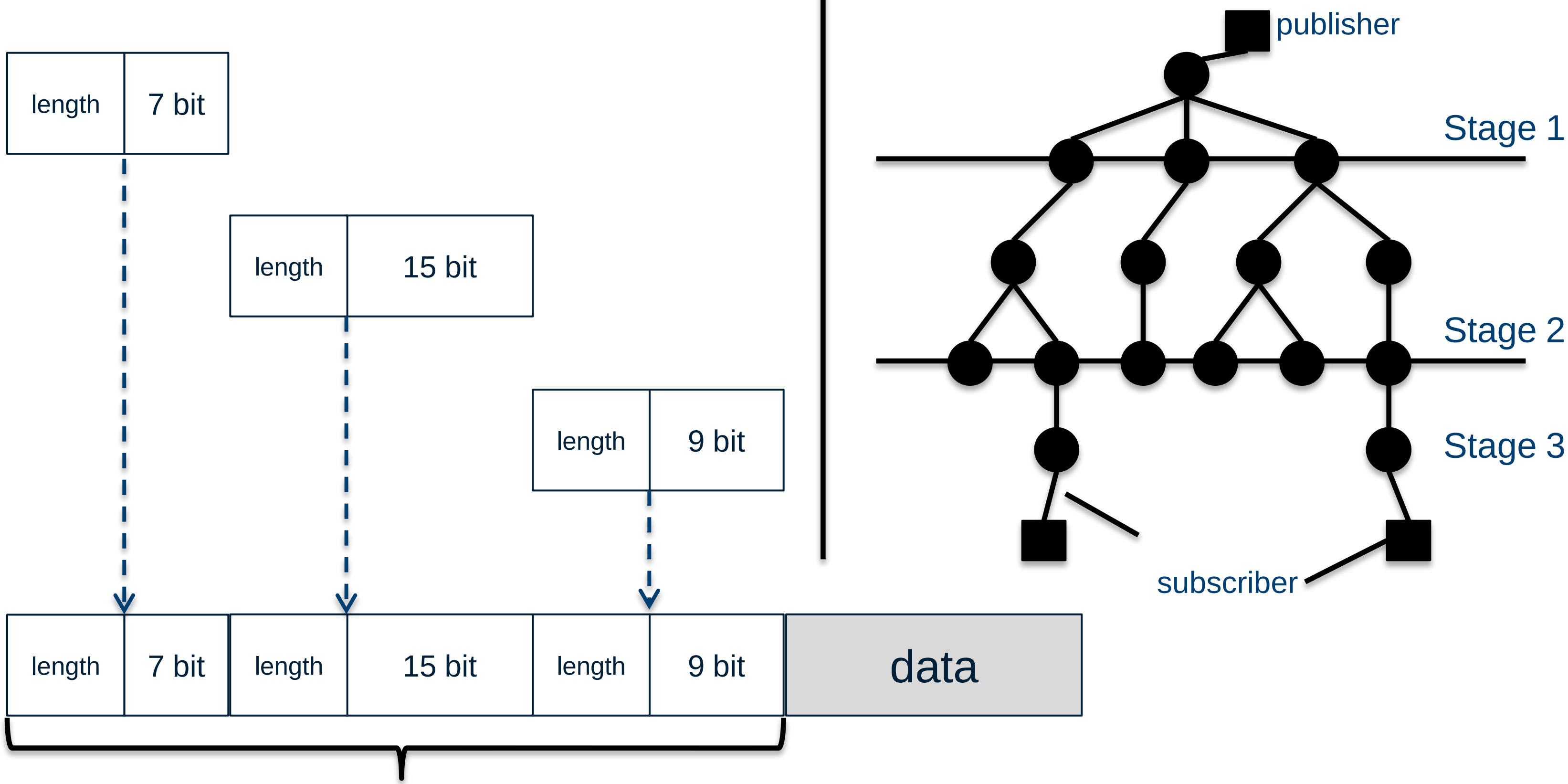


- Perform path computation in AS_1 (delivered from rendezvous results)
- Formulate e2e path from pub in AS_1 to subs in AS_2 and AS_4
- Pros: Addresses both constraints 1 and 2
- Cons: Violates both constraints 3 and 4, i.e., requires knowledge of AS-internal topology and LIPSIN does not scale beyond certain false positive rate threshold

Address Violation of Constraint 4: Multi-stage BF Forwarding

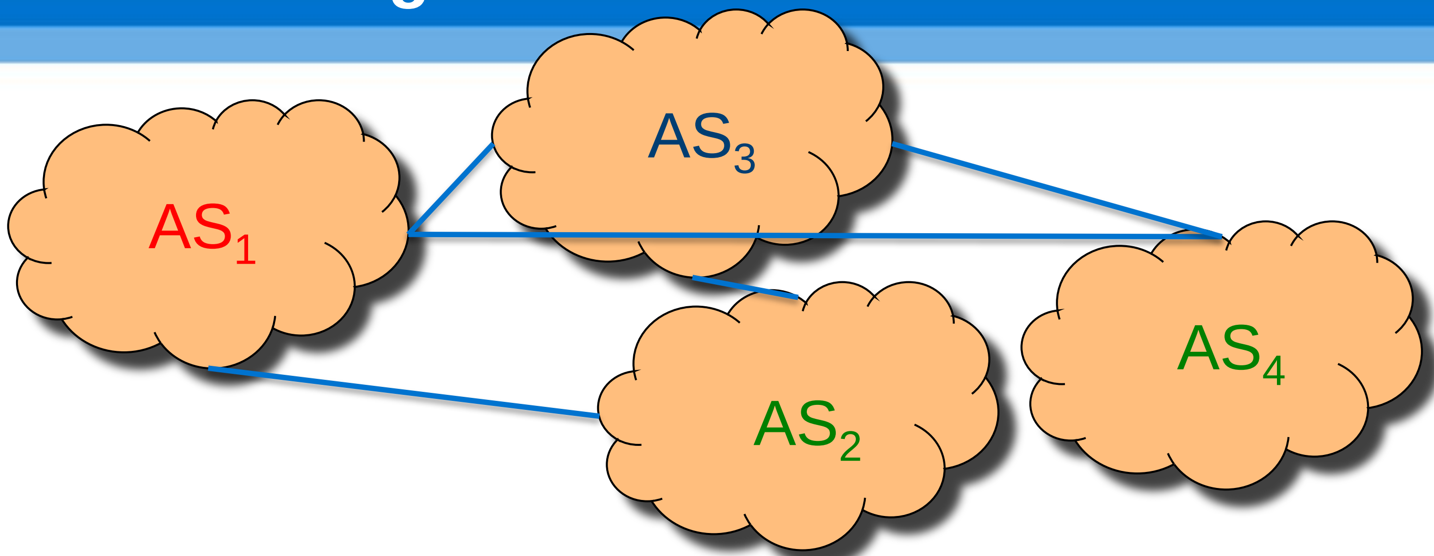
Stage-level forwarding identifiers

Delivery Tree



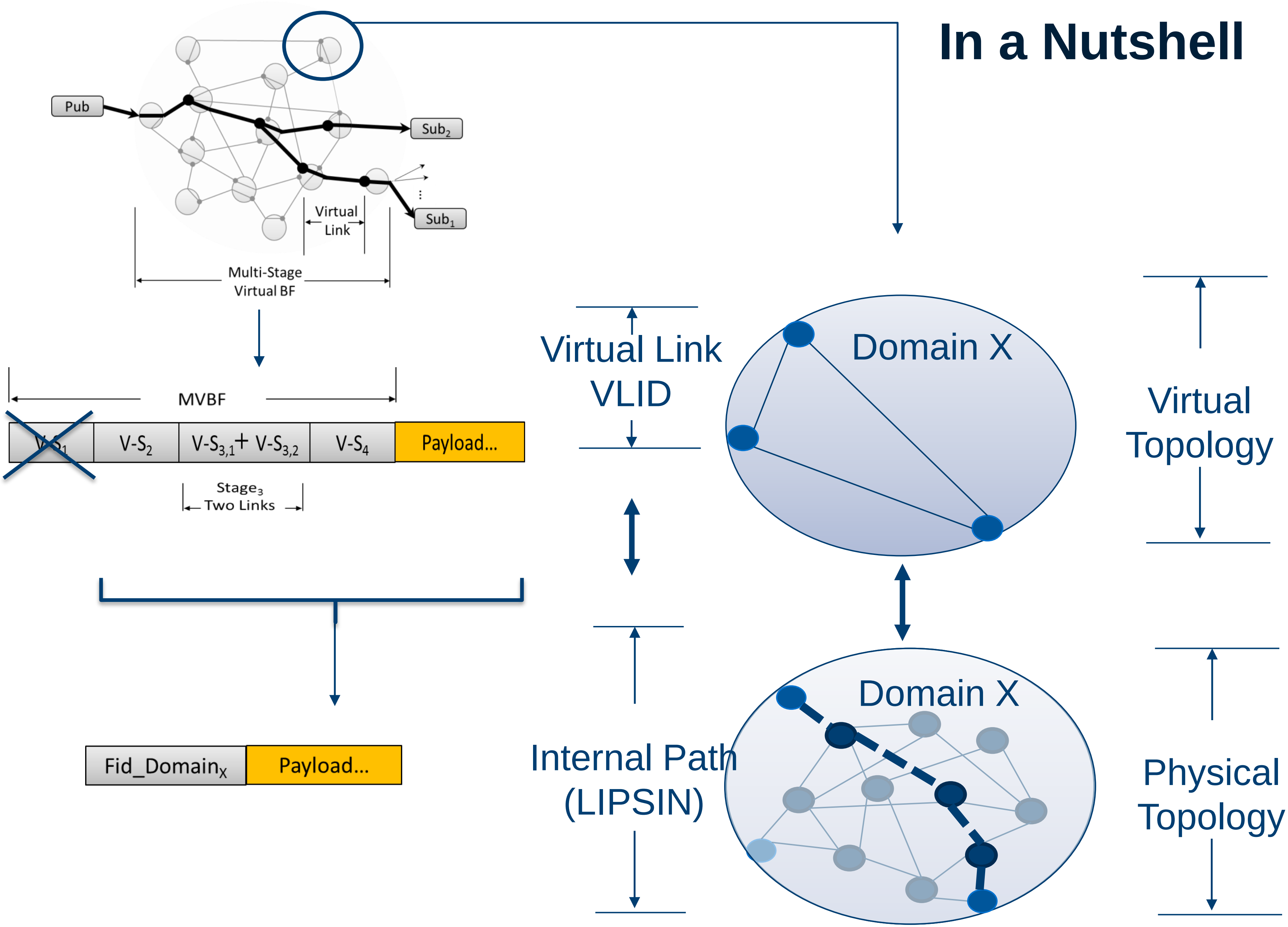
Nodes here are AS-level networks!

Address Violation of Constraint 3 (and Fulfill 1): Pathlet Routing



- Virtualise ingress/egress path (i.e., AS-AS connectivity) through individual virtual link
- Create policy-compliant pathlet from AS_1 to AS_2 and AS_4 as list of virtual links between intermediary ASes
  - Source forwarding compliant with LIPSIN and MS-BF idea
  - Path computation done in topology manager of AS_1
- AS-internal forwarding between ingress and egress done through encapsulation, providing freedom of choice for intra-domain solution

In a Nutshell



Local Reachability: Distributing in Subscriber ASes

- Core Node approach
  - Acts as subscriber to inter-domain information, re-publishing at intra-domain level
  - Use domain-local rendezvous involvement (could optimize through implicit rendezvous strategy) to subscribe
- Optimizations
  - Any domain-local delivery mechanism is supported!
  - Pro-active mode can start setup of core node once local subscribers subscribe -> setup core node in parallel to global rendezvous and path calculation -> likely reduce delay to zero

Gains from This Solution

- Border Node Complexity
  - Limited by AS connectivity rather than Internet connectivity

  -> border routers LESS complex than today's IP-based ones!

- Topology Manager Complexity
  - Similar to today's BGP-based routers (in terms of table size), BUT
    - Tables only needed for path computation, not forwarding -> possible to use cheaper memory
    - Can reduce table size when using non-optimal routing (partial dissemination of inter-domain routes)

Gains from This Solution (2)

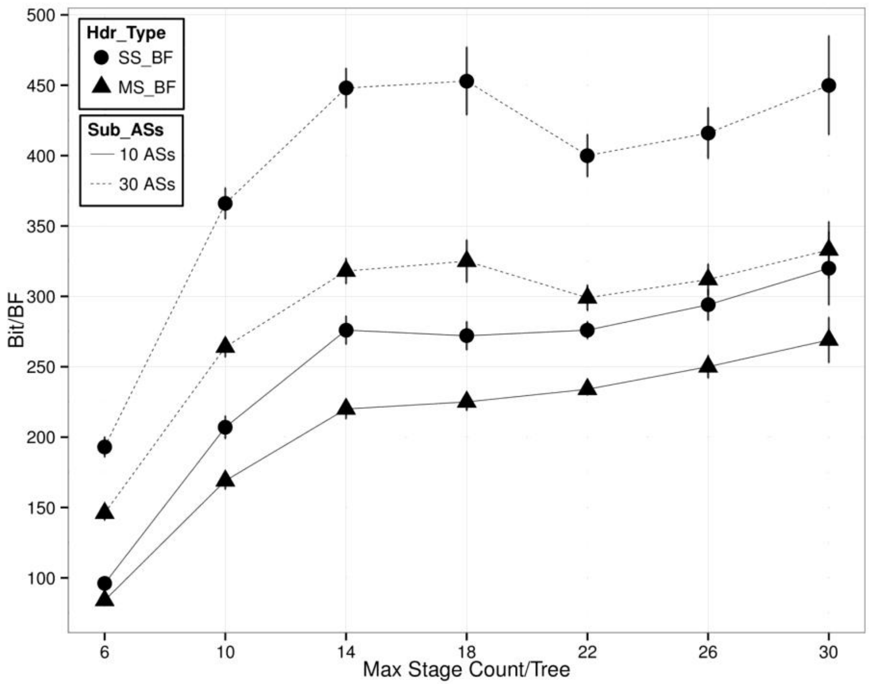
- Policy Compliance
  - Policies incorporated into path computation, not forwarding

  -> more complex contracts envisioned as being possible, executed in topology manager

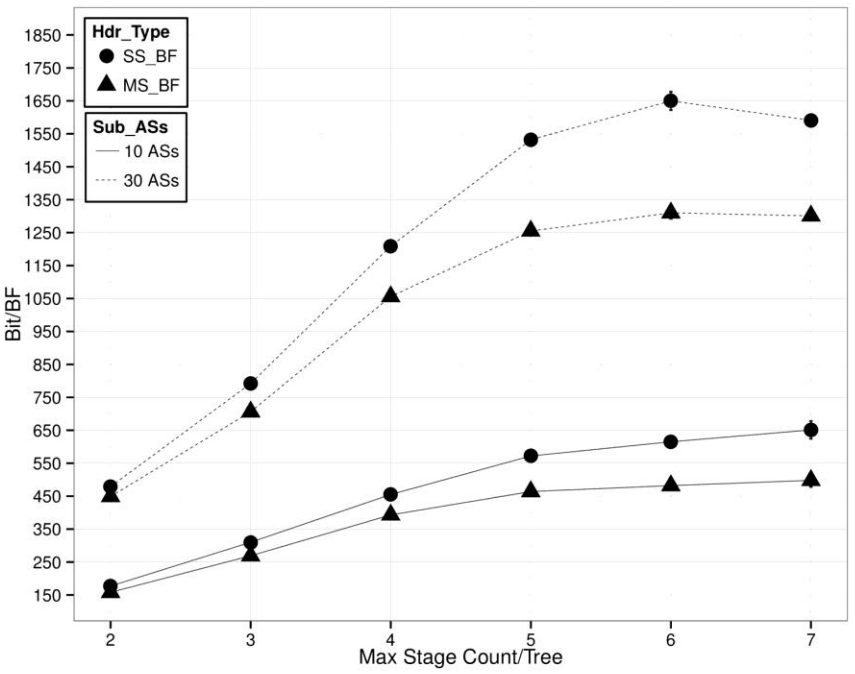
- Forwarding overhead
  - Evaluations show limits of 500 bits header length for large inter-domain routes -> definitely less than IP source routing
  - Inherent and efficient multipoint support

Header Length at Inter-Domain Level

Multicast Header Length for Single & Multi-Stage BF for USCarrier Network with Bound Path Length



Multicast Header Length for Single & Multi-Stage BF for CAIDA Network with Bound Path Length



UNIVERSITY OF CAMBRIDGE

Border Node Complexity: Processing Tables

Assumption: Hash tables used for efficient processing!

-> how large will hash tables need to be?

- Results show mean of 156 and maximum of 507 bits for inter-domain header length
- Set *m* to maximum of 512 for a false-positive free delivery
- Pre-compute the hashes for each virtual link ID and varying *m*.
- Store in CAM with 512 entries for efficient processing

Border Node Complexity: Mapping Tables

Vlink type	Lowest Quartile	Median	Mean	Upper Quartile	Max
Core	17	27	38.6	50	753
Peer	1	2	8.79	4	2097

- Virtual links to core nodes (based on Internet Topology Zoo data)
  - assume co-locating core with forwarding node
  - 75% only require **50** virtual links
- Virtual links to peer ASes (based on CAIDA data)
  - 75% only require **4** virtual links
- **Total (binary!) CAM size** = 512 * (50 + 4) = **27648** entries for 75% of ASes

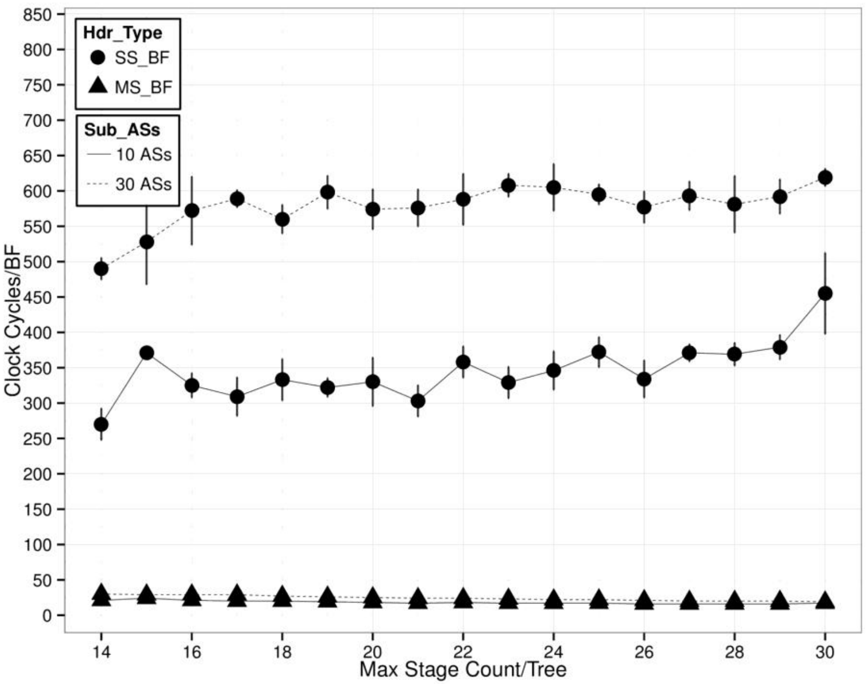
SIGNIFCANTLY less than BGP table sizes for most ASes

Similar for a small number of Ases

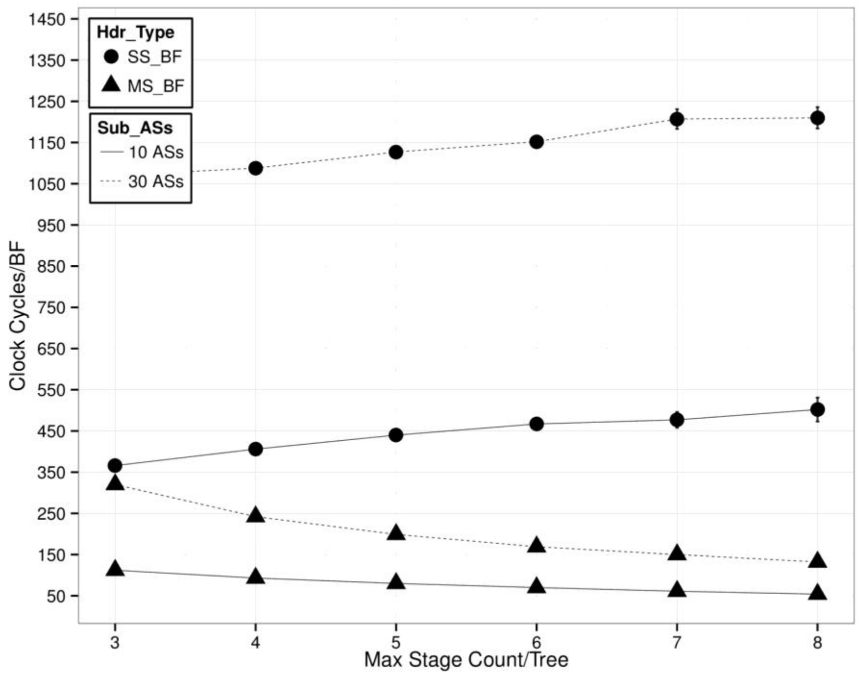
Bounded by own AS connectivity NOT the size of the Internet!

Border Node Complexity: Processing

Processing Delay of Multicast Trees for USCarrier Network



Processing Delay of Multicast Trees for CAIDA Network



Opportunities for Traffic Engineering

- Traffic separation
  - Expose one virtual link for BE, one for QoS
- Resilience
  - Expose more than one virtual link per core/peer relation
- Multi-layer environments
  - Realise different virtual link mappings
- Caching
  - Fill AS-internal caches at core nodes

Open for Future Work

- Policy expression
  - How? Which ones? How to enforce?
- Deployment
  - How to avoid the IP multicast fate?
- Implementation
  - Can be done through our Blackadder platform

Disseminating Content using Digital Fountains

“The flow is the enemy” – David Oran (Cisco)

Introduction

- … from a well-defined endpoint-to-endpoint model
- … to a loosely coupled multipoint one
- Information can be contributed by many sources for many receivers
- Interested nodes come and go…

Digital Fountains

Moving away from strict flow-based approaches

- Fountain coding to encode content
- Self-contained, cacheable encoded symbols
  - Embedded in the (information centric) identification scheme
- Decoupling dissemination from the management
- Many (including caches)-to-many communication
- Multi-path support

Digital Fountains

- Content Publisher:
  - Initial content is fragmented (constant and known size)
  - A very large number of encoded symbols is created as a binary combination of some fragments (XOR)
  - The **art of fountain coding*** is to select the *degree* of each symbol!!
    - How many packets to XOR in the symbol?
    - *Neghbours' set (uniformly selected)*

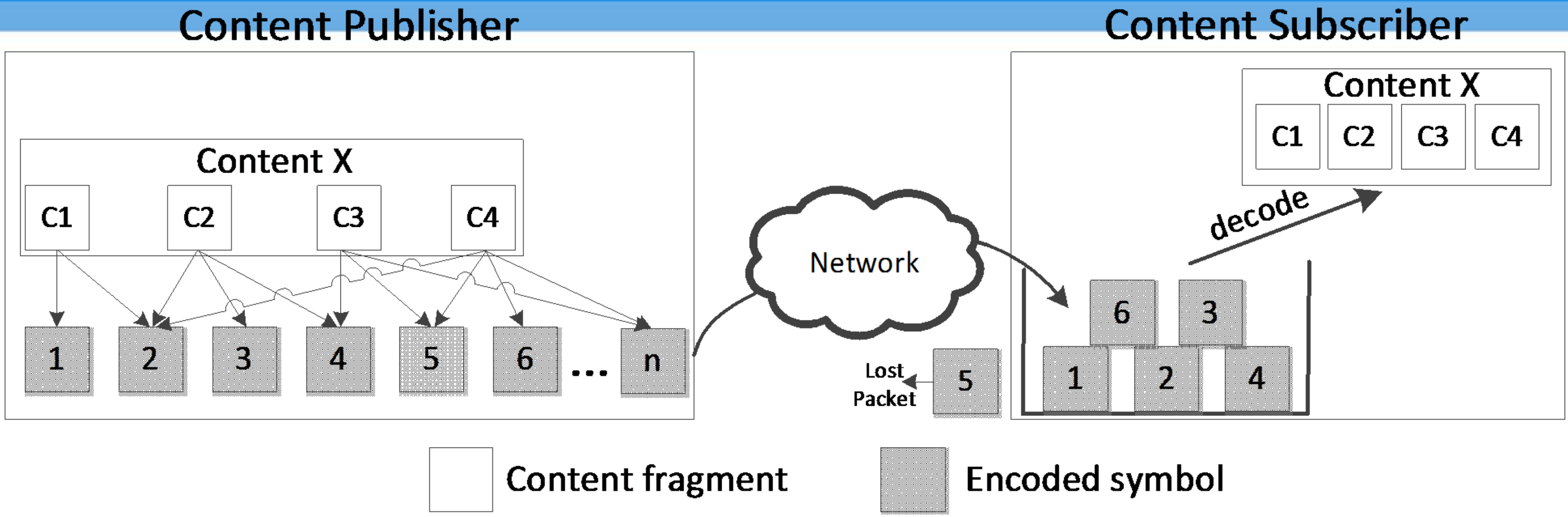
*: read "intellectual property"

Digital Fountains

- Content Subscriber:
  - If a symbol has *degree 1*, it is decoded…
  - Using this content fragment, the degree of all symbols that contain it is decreased by 1 (XOR)

Some coding related information needs to be communicated!

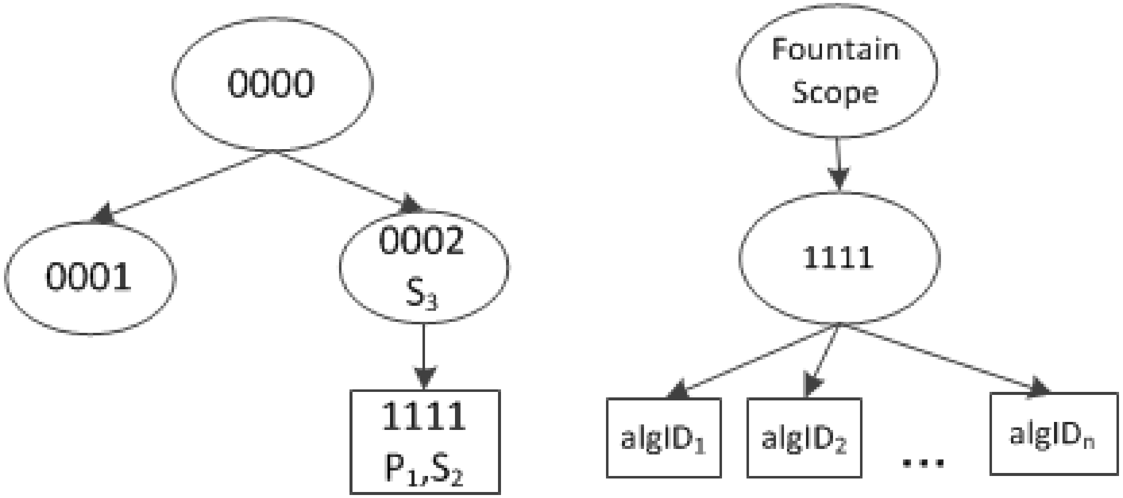
Digital Fountains



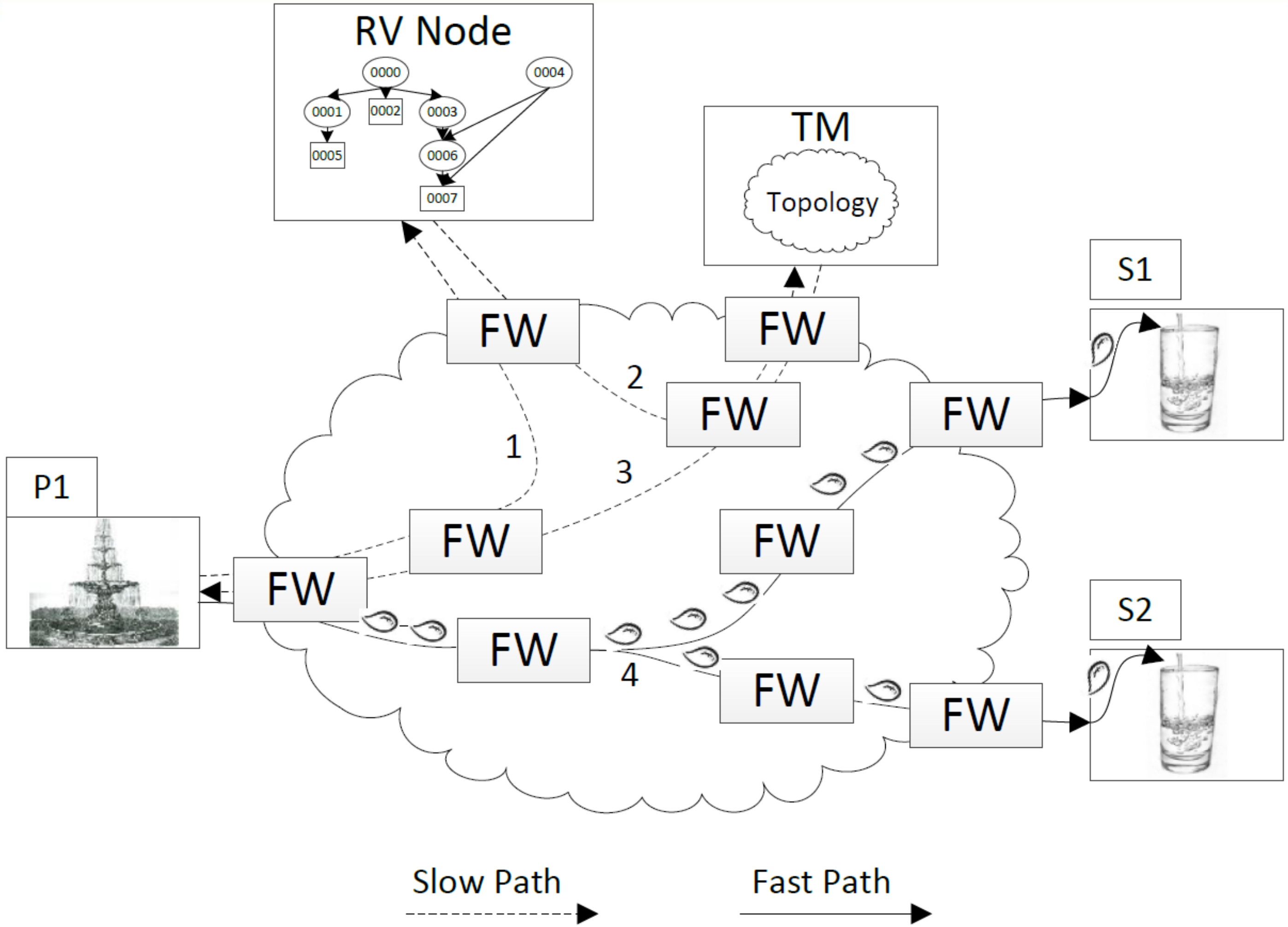
- Lost packets affect only the efficiency of dissemination
- No retransmissions are required – no feedback channel
- A subscriber needs a number of symbols (slightly larger than the number of fragments) to decode the content

Labelling Encoded Symbols

- For some content identified as /a/b/c encoded symbols are identified as $/x/a/b/c/algID_n$ or $/x/algc/algID_n$ or just $/x/c/algID_n$
- $algID_n$ contains enough information for subscribers to decode the symbol
- /x/y is a scope dedicated for publishing encoded symbols only
  - …known by



Basic Operation

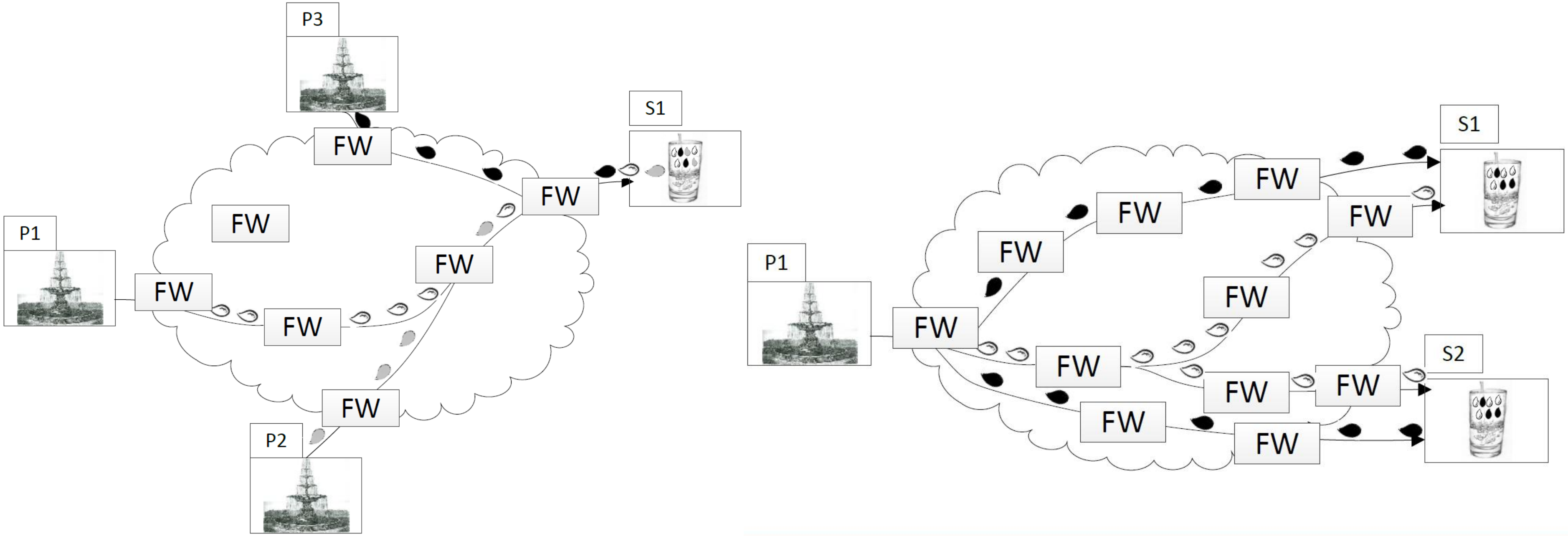


Asynchronous Operations

- A subscriber can join a fountain at any time
- The content is decoded as long as the required number of symbols is received
- The fountain stops only when no subscribers exist
- Digital fountains can produce a very large number of unique symbols

Multi-Publisher & Multi-Path/Source

- At a given time one or more network nodes (incl. network caches) may store…
  - the decoded version of a content
  - encoded symbols of a content



In-network Storage

- In-network nodes can subscribe to parts of the information structure (implicitly)
- Special Link Identifiers that “point” to a caching component can be included in a LIPSIN identifier by the TM
  - Feeding the network with encoded symbols when/where needed
  - A separate control point gives power over the caching strategies
  - …instead of caching everything everywhere
- CDN-like: A node decodes the content and becomes a publisher (RV is notified)
- Opportunistic: Nodes advertise an item without decoding it

Application: Personalised Media

Content but highly personalised!

Slides thanks for Ben Tagger and Stuart Porter (CTVC)

Media Usage in the Current Internet…

YouTube

- 4 B hours of videos per month.
- 1 T views in 2011.
- 25% of these from Mobile devices.

BBC

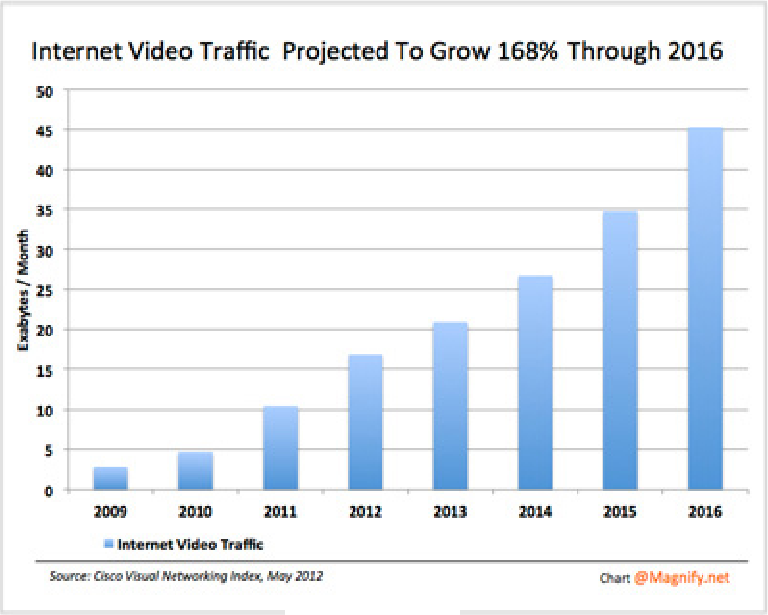
- 190m requests per month.
- +25% from last year (mobile +95%).

$$$

- AOL Video Ad Revenue
- $10m (2010) to $100m (2012)

Future - 2016

- 1.2m mins of video per sec.
- 86% of global traffic.
- Online Video Industry - $28b

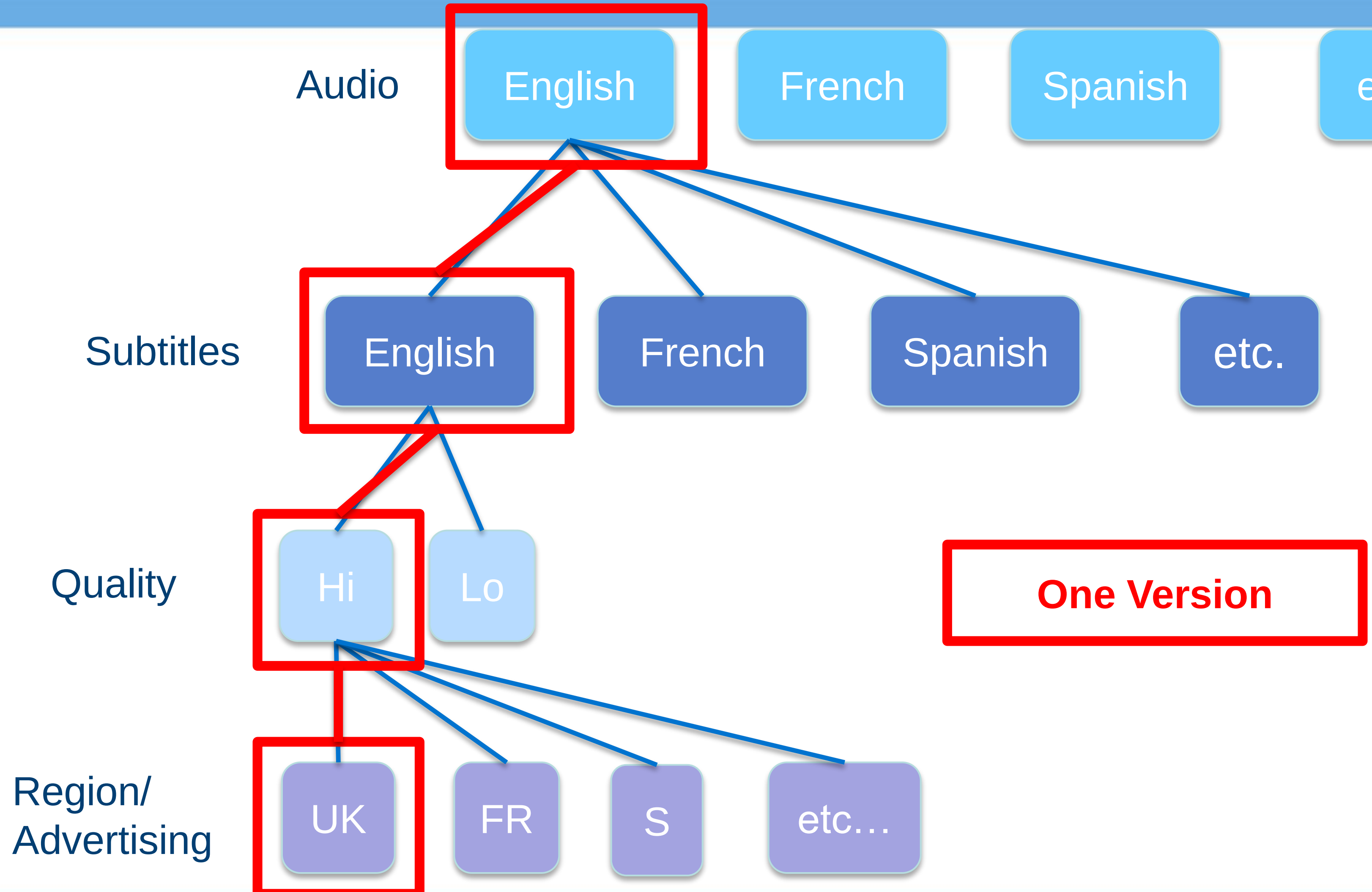


CISCO

UNIVERSITY OF CAMBRIDGE

TecVis

The problem of combinatorial complexity



What is the problem?

- Content production companies are producing more media and more versions
  - Plethora of channels
  - New international markets
- Different channels and markets have different requirements
  - Length & ad breaks
  - Tussles between content & local restrictions
  - Delivered Format
  - Subtitles, translations and aural description files

What is the problem?

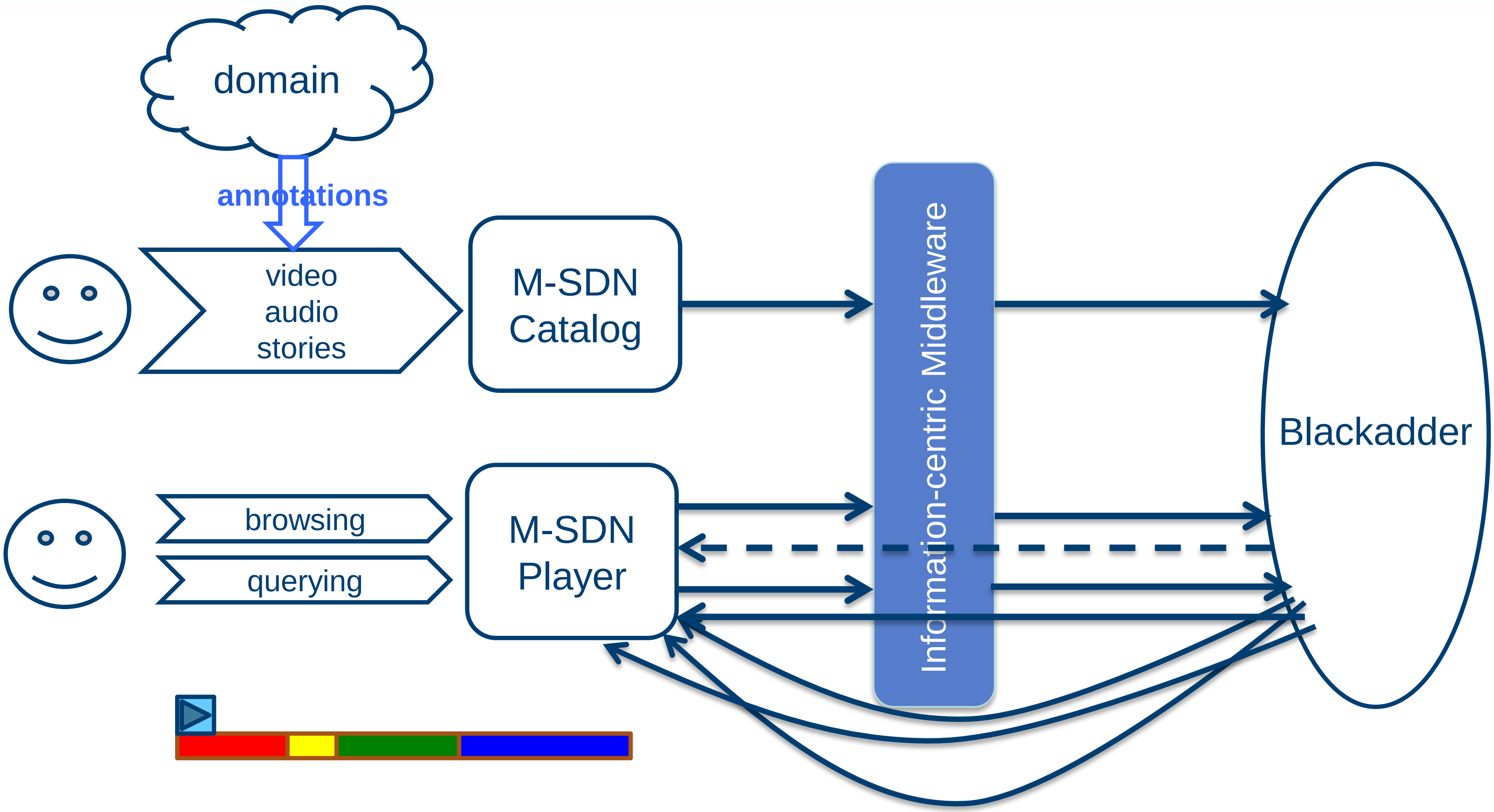
- Expensive
  - Extra days of work in the edit suite
  - Multiple viewing copies created in multiple formats
  - Duplication costs & international transportation
  - Low res (mobile) High res (HDTV)
- Currently network cannot identify that large parts of different files are identical
  - Caching not possible to save resources and time
  - Hard to pull parts from separate repositories

Is it important?

- Yes
  - Broadcasters are losing money
  - Therefore they pay less for content
  - Content producers who still have high overheads
  - Producers need to find ways to cut costs without cutting quality, quantity or staff.
- So a digital rights management solution where an overall storyline can be adapted based on the dynamically changing rights of a user would be beneficial.

Current practice

The M-SDN: Overview



Why is this Difficult in the Current Internet?

- Fine-grained access to Networked Media
  - Access control, parental control and ethical constraints.
- Serving personalised, democratised content.
- Logistics of version storage leading to version management nightmare.
- Current: Delivery of monolithic blobs of media.
- **M-SDN: Delivery of personalised, distributed and dynamic media experiences.**

Information-Centric Middleware

Currently…

- ICN places emphasis on information at the Network layer,
- Blackadder uses a pub-sub architecture,
- any applications must operate directly on this layer.

We want…

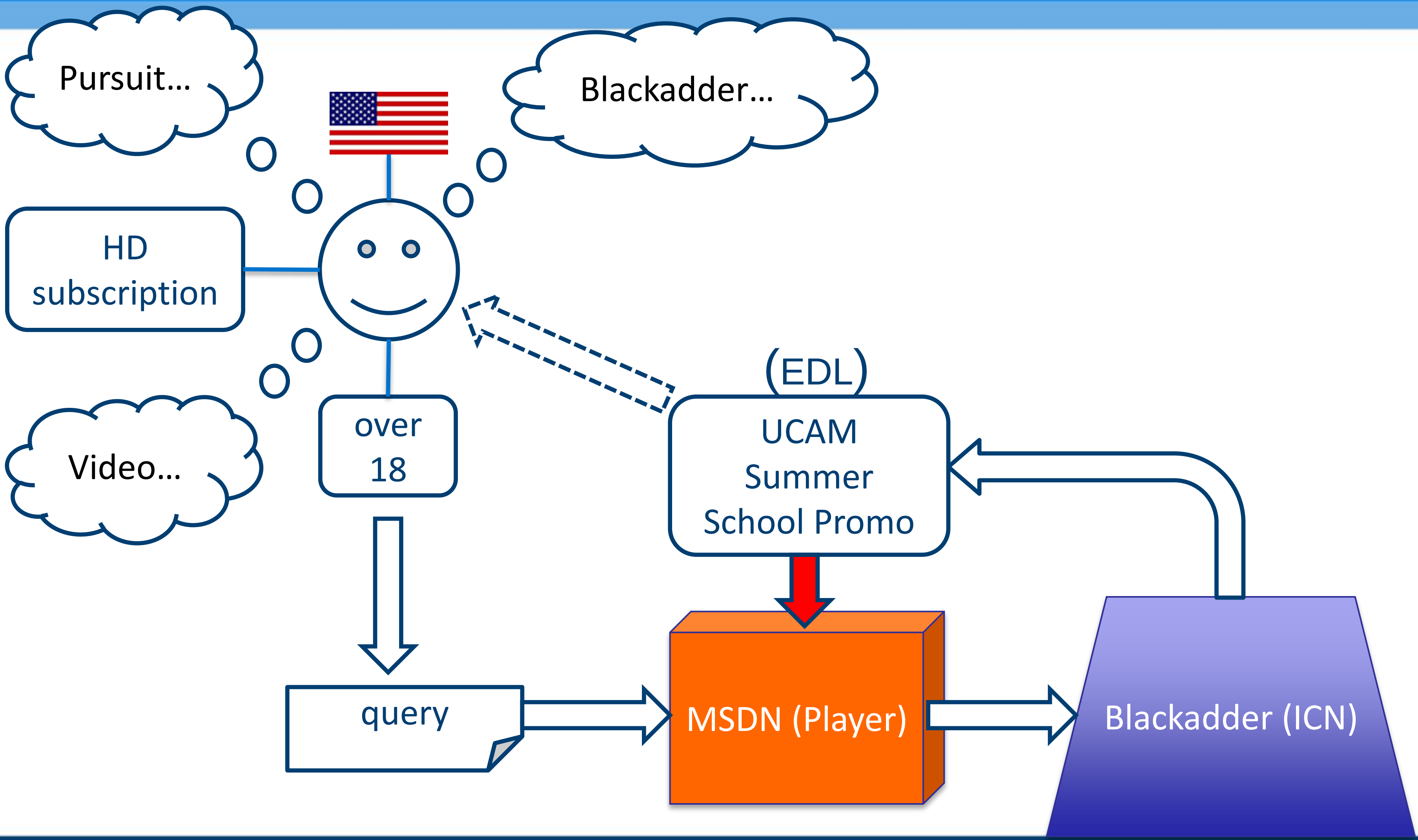
- the emphasis on information to continue at the application layer.
- to provide natural abstractions that facilitate ICN use.
- So, we need a middleware.

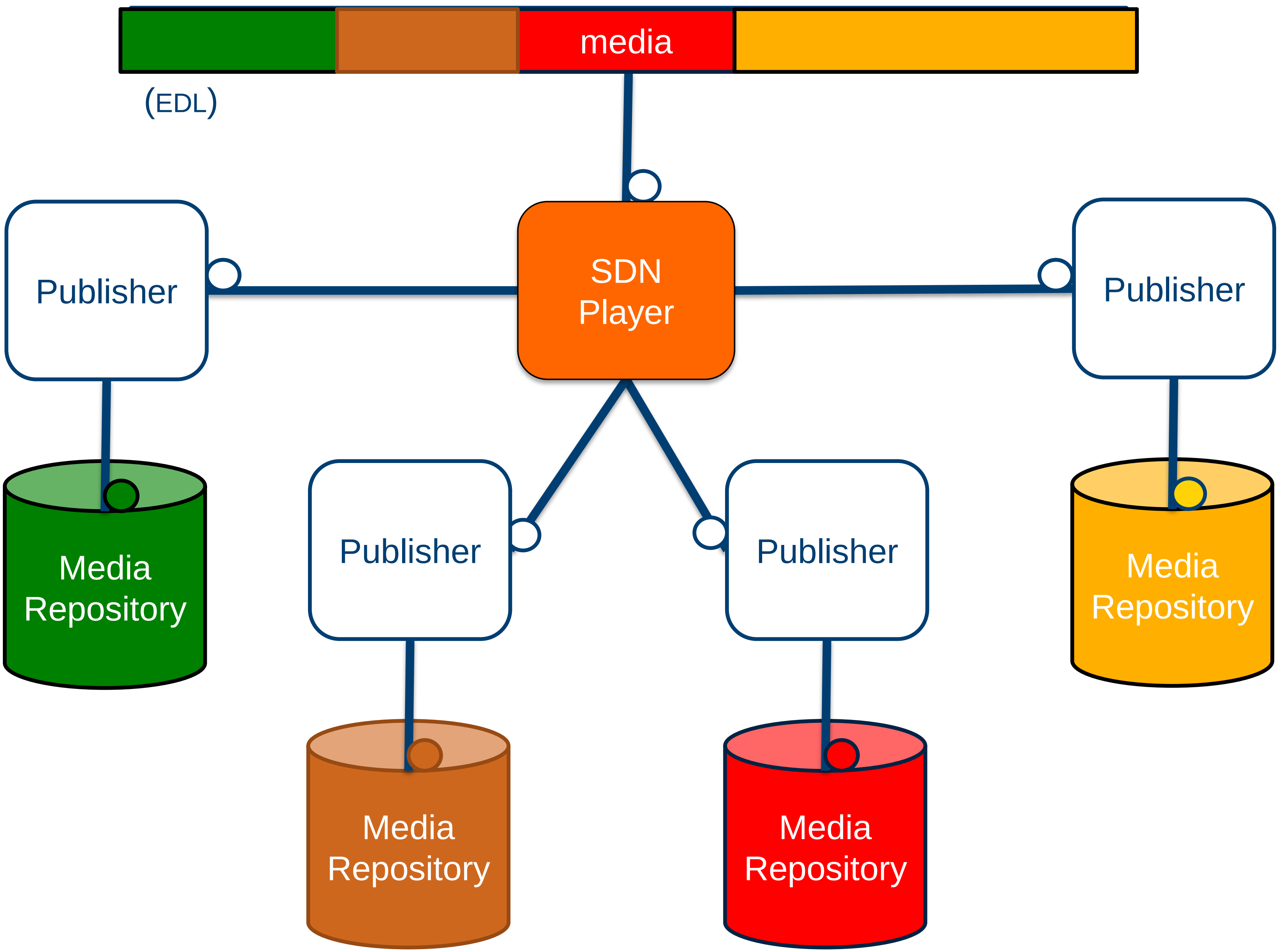
Features of the Middleware

Our middleware…

- extends native ICN metamodelling, adding expressivity,
- enforces a consistent, satisfiable network metamodel,
- leverages these features to enable (non-native) network mechanisms, such as
  - browsing,
  - querying/searching,
  - distributed querying.

The M-SDN: how do we get media out?




media
(EDL)
SDN
Player
Publisher
Publisher
Publisher
Publisher
Media
Repository
Media
Repository
Media
Repository
Media
Repository

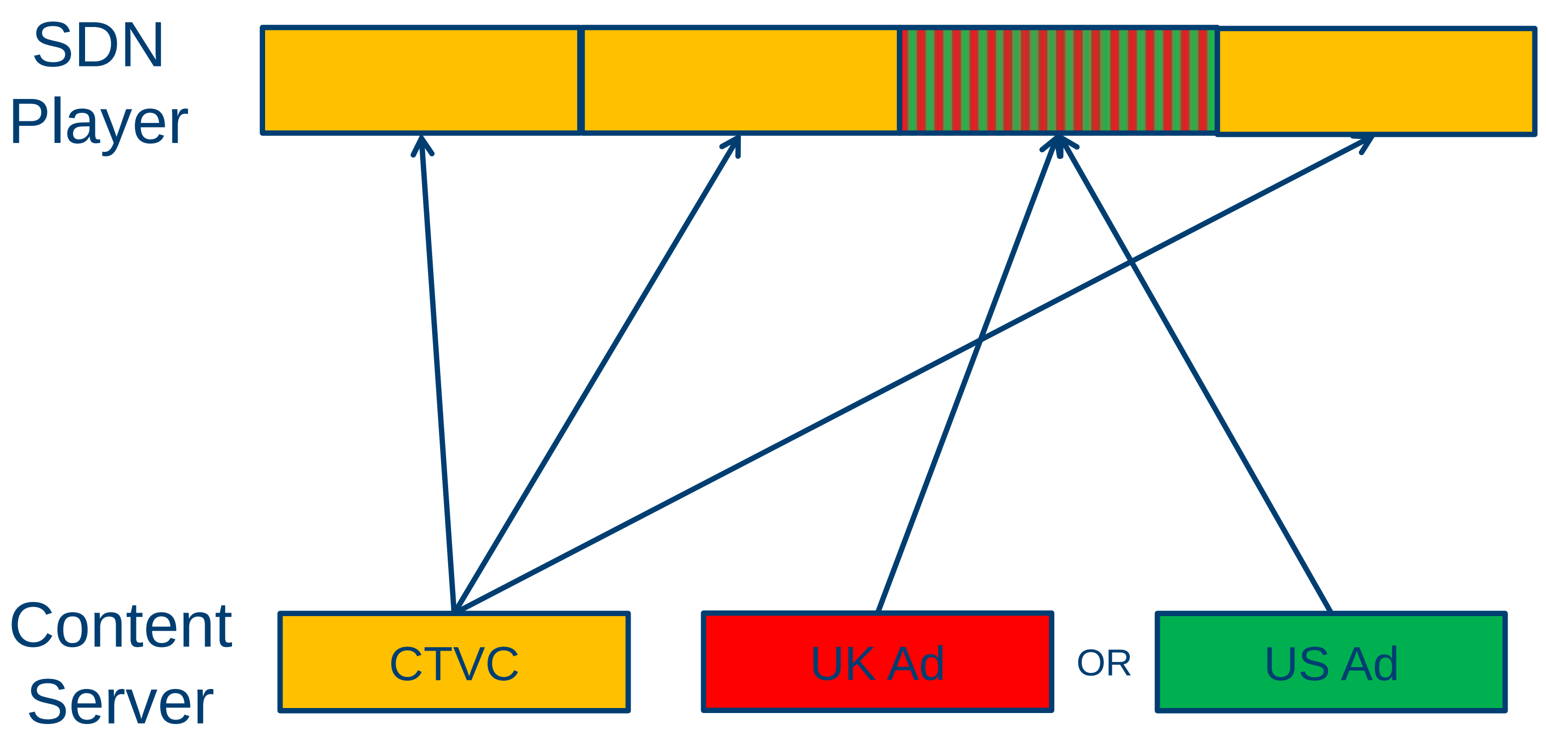
Demonstration – Personalised media!

The differences between 2 EDLs for the same video subscribed to from the USA and the UK

```
TITLE: EDLTEST1
FCM: NON-DROP FRAME

001  D1T1C3A  AA/V  C        00:00:51:00 00:01:01:00 00:00:00:00 00:00:07:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#D1T1C3-GVs-b

002  D1T1C3A  AA/V  C        00:00:07:00 00:00:15:00 00:00:07:00 00:00:14:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#D1T1C3-GVs-b

003  D3T8C1A  AA/V  C        00:00:00:00 00:00:09:00 00:00:31:00 00:00:37:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#ADVERTDanyaUK

004  D1T1C3A  AA/V  C        00:00:21:00 00:00:30:00 00:00:14:00 00:00:21:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#D1T1C3-GVs-b-1

005  D1T1C3A  AA/V  C        00:00:34:00 00:00:46:00 00:00:21:00 00:00:31:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#D1T1C3-GVs-b-1

006  D3T8C1A  AA/V  C        00:00:35:50 00:00:45:50 00:00:31:00 00:00:37:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#D3T8C1-GVs-a

007  D3T8C1A  AA/V  C        00:00:43:50 00:00:51:50 00:00:37:00 00:00:43:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#D3T8C1-GVs-a

008  D3T8C1A  AA/V  C        00:00:34:00 00:00:45:00 00:00:43:00 00:00:51:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#D3T8C1-GVs-b

009  D3T8C1A  AA/V  C        00:00:08:00 00:00:25:00 00:00:51:00 00:01:09:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#D3T8C1-GVs-b-1

010  D3T8C1A  AA/V  C        00:00:22:00 00:00:48:00 00:01:09:00 00:01:32:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#D3T8C1-GVs-b-1
```

```
TITLE: EDLTEST1
FCM: NON-DROP FRAME

001  D1T1C3A  AA/V  C        00:00:51:00 00:01:01:00 00:00:00:00 00:00:07:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#D1T1C3-GVs-b

002  D1T1C3A  AA/V  C        00:00:07:00 00:00:15:00 00:00:07:00 00:00:14:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#D1T1C3-GVs-b

003  D3T8C1A  AA/V  C        00:00:00:00 00:00:09:00 00:00:31:00 00:00:37:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#ADVERTAsimUSA

004  D1T1C3A  AA/V  C        00:00:21:00 00:00:30:00 00:00:14:00 00:00:21:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#D1T1C3-GVs-b-1

005  D1T1C3A  AA/V  C        00:00:34:00 00:00:46:00 00:00:21:00 00:00:31:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#D1T1C3-GVs-b-1

006  D3T8C1A  AA/V  C        00:00:35:50 00:00:45:50 00:00:31:00 00:00:37:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#D3T8C1-GVs-a

007  D3T8C1A  AA/V  C        00:00:43:50 00:00:51:50 00:00:37:00 00:00:43:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#D3T8C1-GVs-a

008  D3T8C1A  AA/V  C        00:00:34:00 00:00:45:00 00:00:43:00 00:00:51:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#D3T8C1-GVs-b

009  D3T8C1A  AA/V  C        00:00:08:00 00:00:25:00 00:00:51:00 00:01:09:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#D3T8C1-GVs-b-1

010  D3T8C1A  AA/V  C        00:00:22:00 00:00:48:00 00:01:09:00 00:01:32:00
* FROM CLIP NAME:  http://localhost:8080/owl/review_catalog.owl#D3T8C1-GVs-b-1
```

Demonstration – Personalised media!



What is the Impact?

- ICN allows increased competition
- Embedding M-SDN in ICN architecture means user can request differentiated services
  - Enables network providers to optimize delivery according to implicit and explicit preferences of subscribers
  - Potentially leading to
    - Fairer pricing strategies
    - New market mechanisms for content delivery
    - Move away from flat pricing structures

What is the Impact?

- M-SDN can also allow changes to current models
  - Exponential growth of internet traffic – esp. video
  - Investment without direct positive effect on revenue
  - Content providers are increasing profits - at the expense of ISP's investment in capacity

What is the Impact?

- New SLAs possible which allow per item charges for uploads and downloads
  - Content creators pay ISPs' to upload items
  - In turn they are paid by consumers who download items
  - "Visiblity" across the network achieved through M-SDN and ICN could enable money flow from user to provider
  - Users could also monitor SLA's on per item basis

What is the Impact?

- BEWARE
  - Possible adverse effect on privacy
    - Providers could also monitor users on per item basis
  - Users may not take kindly to per item charges
    - Whilst they will agree that it is fairer, users like to know what they are paying each month, and who they are paying it to.

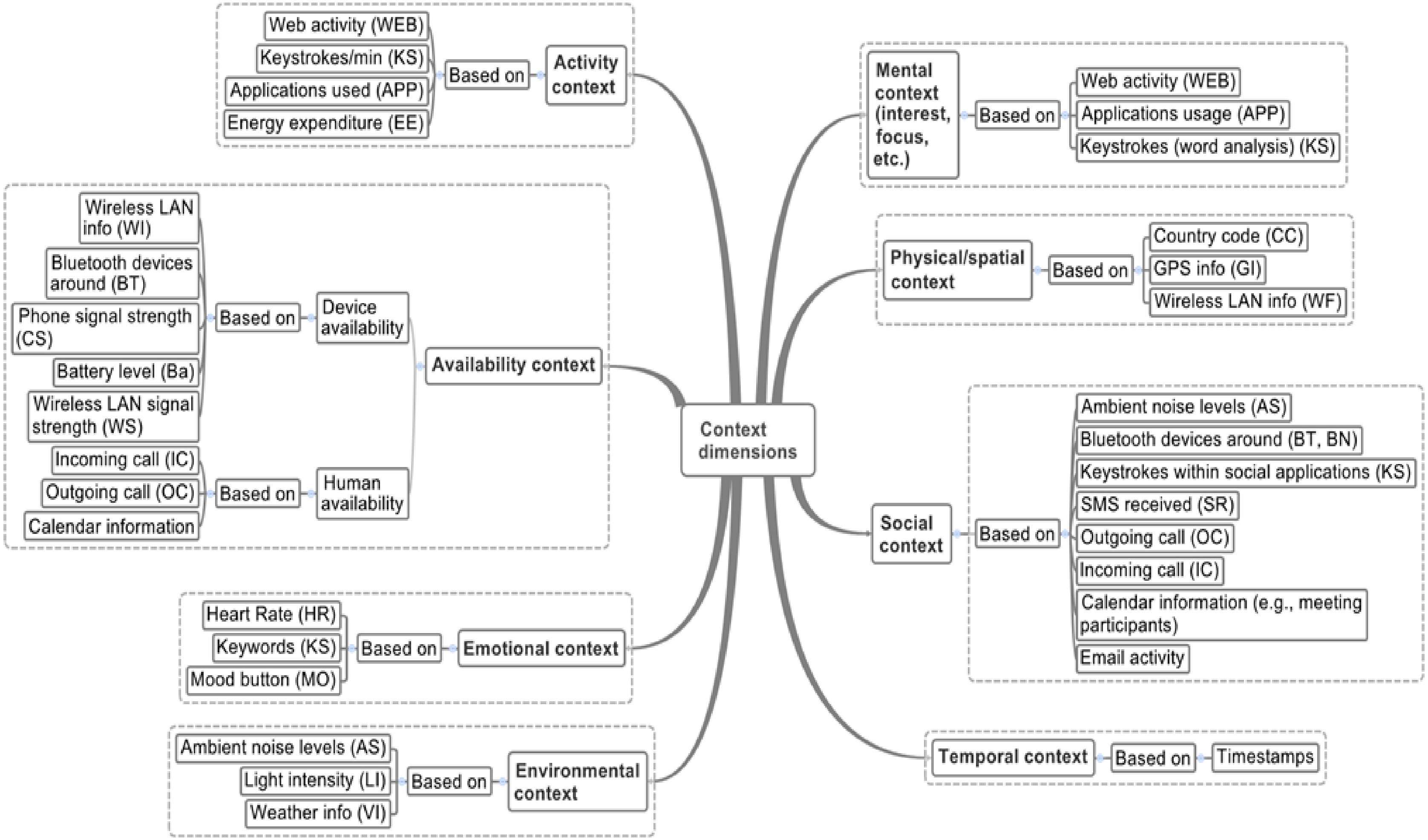
Application: Storytelling

Content but highly personalised!

What is the Problem?

- Individuals recording personal information from their smartphones, laptops, …
- Information is shared at the level of hard as well as aggregated data
- A narrative approach is used to visualise the data to individuals
  - …and share again with your social circles
- Scenarios
  - Long-term conditions
  - Self-awareness
  - Social activism

Plethora of Information



Challenges for ICN

- Individual publishers move about
- Lots of individual (small) data items
  - Around 150k of objects per daily recording at about 20 bytes size

Questions:

- What would happen if such data was published in an ICN?
- What approach makes the data best available?
- What is the right governance approach

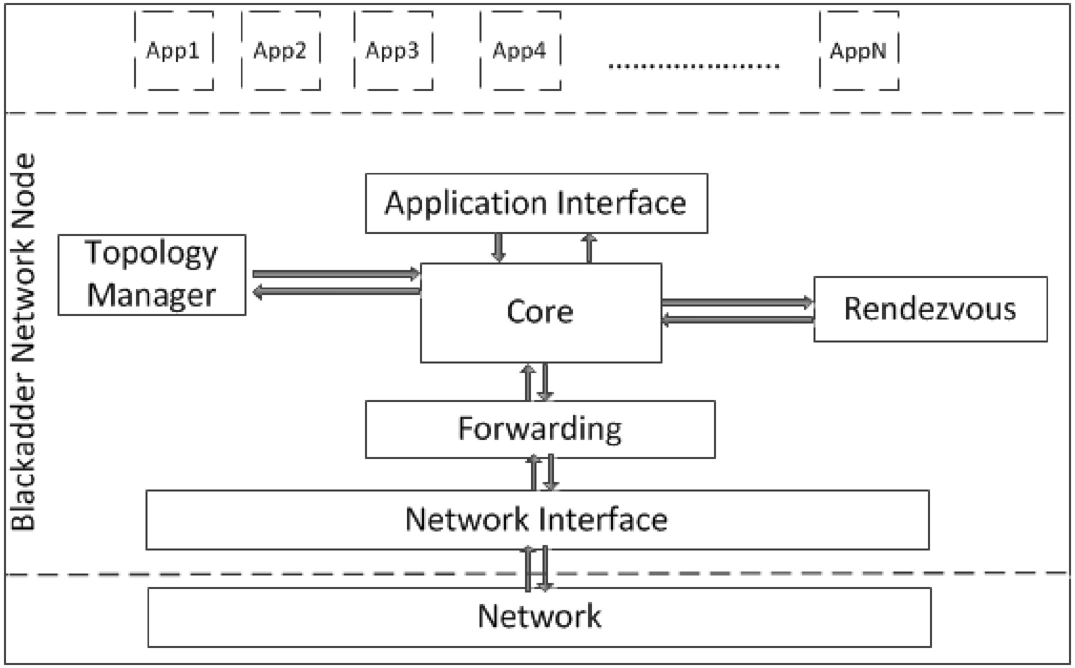
First Glimpse

- Recording application
  - **AIRS** available in Google PlayStore (search for “AIRS Tecvis”)
  - Not yet a pure ICN app (TCP-based pub/sub)
- Visualisation application
  - **Storica** available in Google PlayStore (search for “Storica Tecvis”)
  - Shared over current Internet
- Future work: move towards pure ICN to stress test ICN approaches

Prototype, Deployment & Some Results

Making it work and run - where have we gotten to?

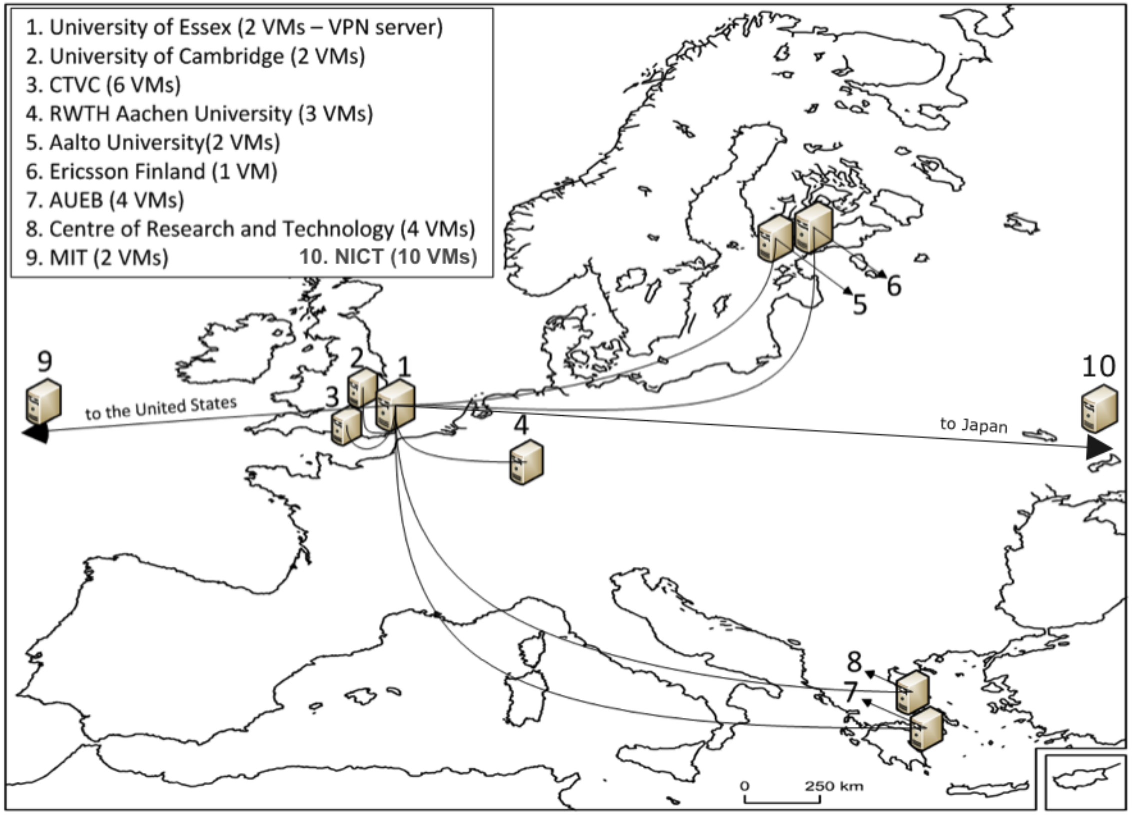
Our Prototype: Blackadder



- Implements design tenets
- Based on **Click** router platform (*)
  - Easy user/kernel space support
  - Easy porting onto other OSes
  - Easy plugging into ns-3
- Available at https://github.com/fp7-pursuit/blackadder
- Domain-local throughput reaches 1GB/s

(*) REF: E. Kohler, R. Morris, B. Chen, J. Jannotti, F. Kaashoek. The click modular router. ACM Trans. Comput. Syst. 18, 3 (August 2000), 263-297.

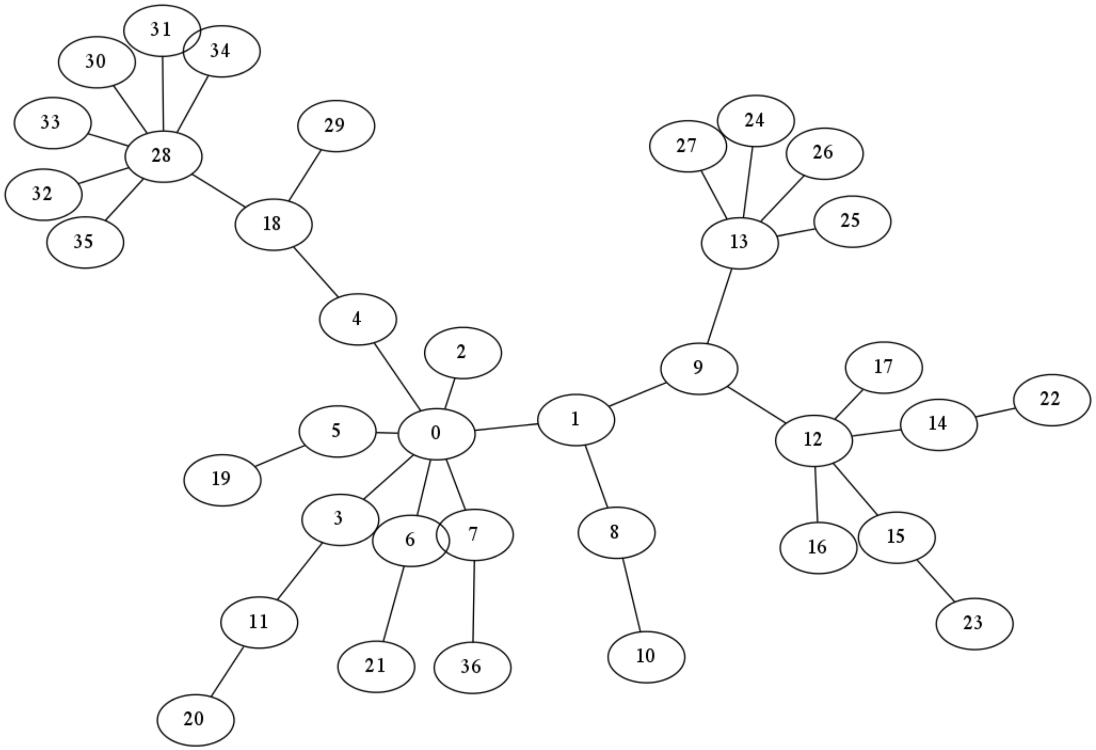
Our PURSUIT Test Bed



- 10 international sites
  - 26 machines with +40 on-demand ones
- Same prototype platform
- More nodes can be plugged in as needed
- OpenVPN-based connections between sites, via UESSEX
- Overlay topologies, different configs possible

Deployment

- Automated deployment
- Deployment tool
  - Topology config → node configurations
  - Address discovery, copying, starting
- Software installation and configuration scripts
  - E.g. dependencies, Blackadder
- For PURSUIT testbed, PlanetLab, other test networks



Example Overlay Topology

PURSUIT Testbed Usage

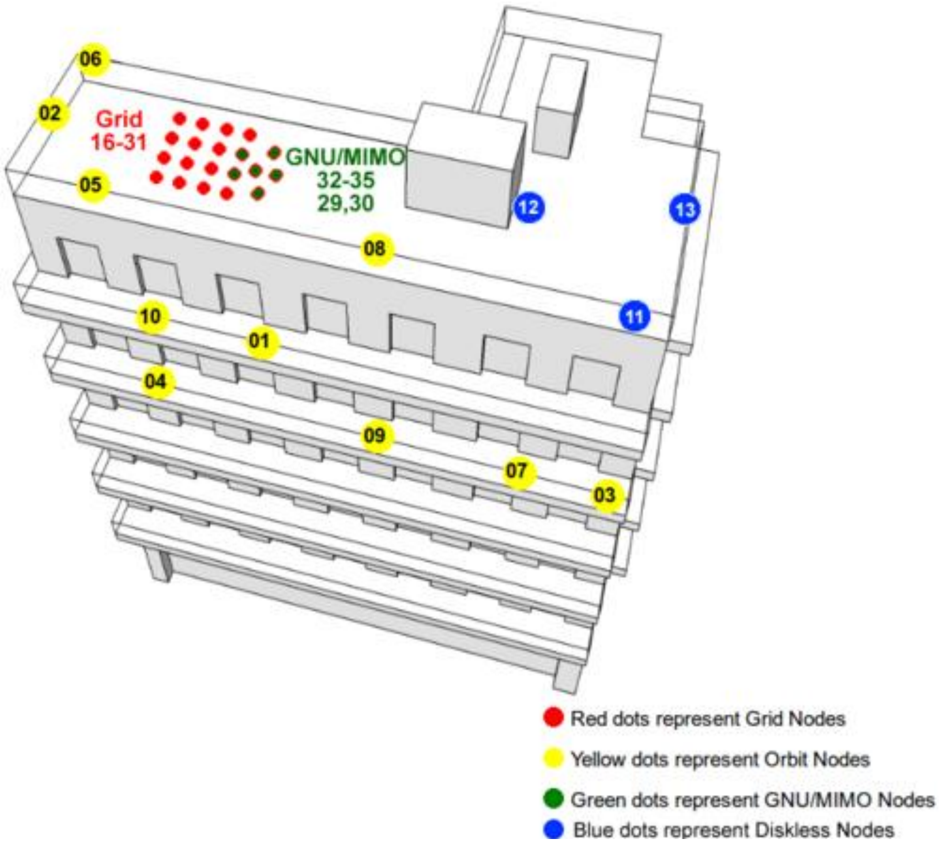
- Testing
  - E.g. prototype, components, applications
  - E.g., when developing new features or improvements, before integrating code into the trunk, before code releases
- Evaluation
- Engagement
  - Trying out the prototype
- Dissemination
  - Demonstrations

Using PlanetLab

- 50 – 200 virtual nodes, custom overlay topologies
- Testing of the prototype
  - Bigger environment, various issues exposed
  - Testing before releases
- Evaluation
  - Scalability (to the extent possible)
- Supported in the deployment tool

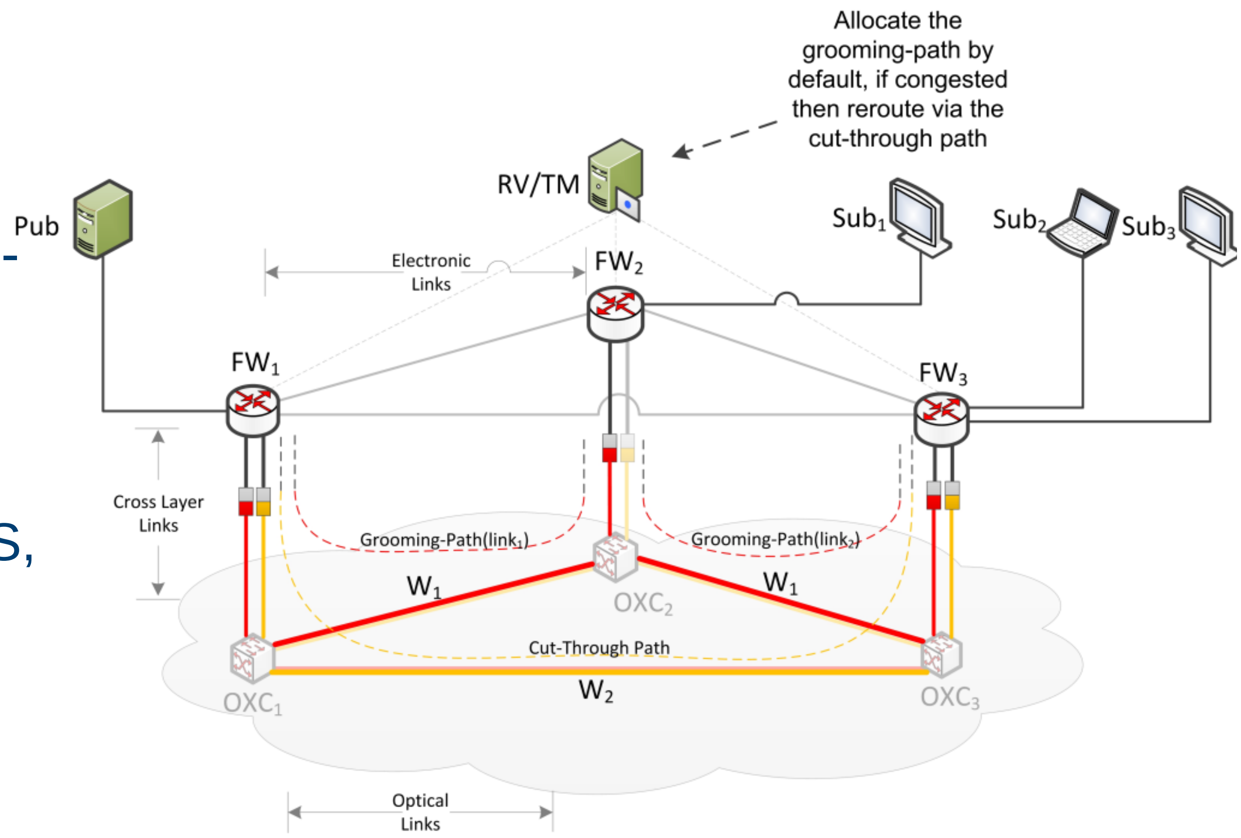
Nitos Testbed @ CERTH

- 20 high-performance nodes
- OpenFlow switches, Gigabit Ethernet links
- Also wireless
- Users reserve physical resources (not VMs)
- Accessible with PlanetLab credentials
- High-speed and SDN experiments

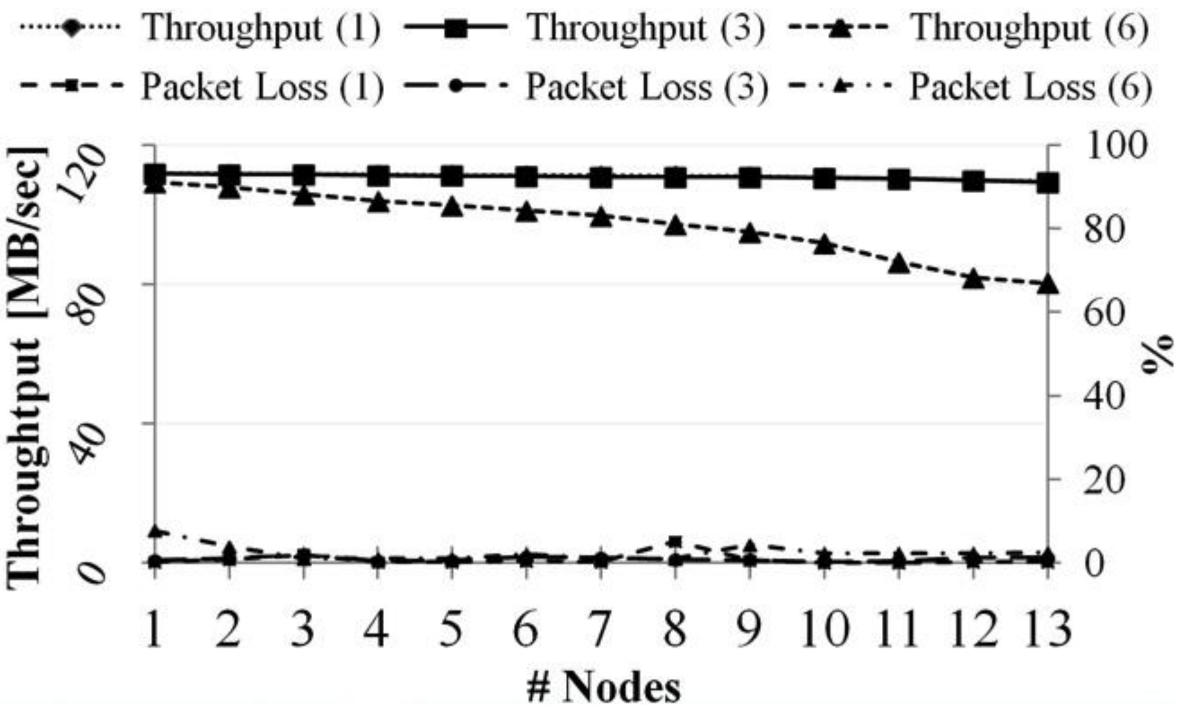


Multilayer Testbed @ Essex University

- Electronic packet-switched layer
- Fast optical wavelength-switched network
- Cross-layer links
- Traffic engineering, QoS, resiliency



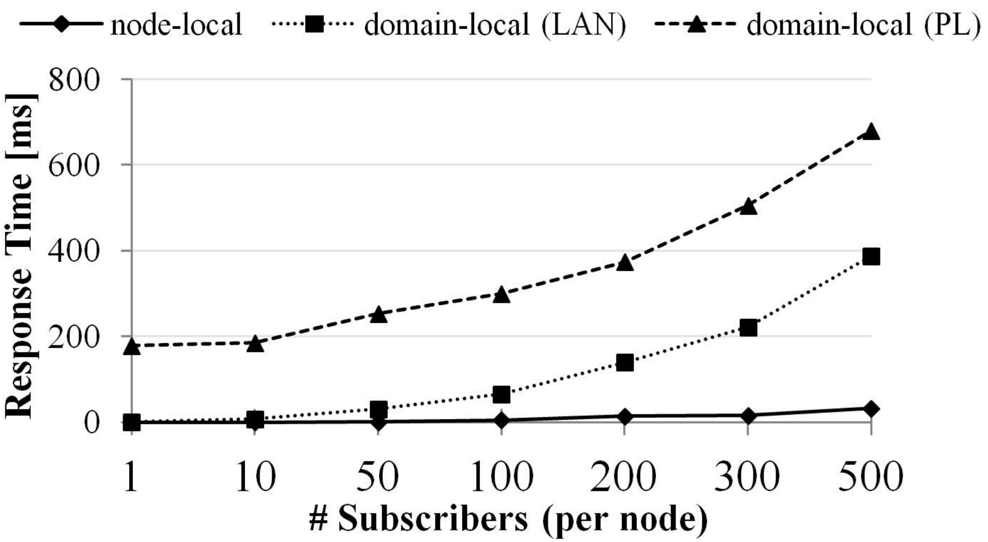
Experimental Evaluation: Fast Path



Forwarding efficiency

- 15 in a chain
- Multicasting (when nodes is sub)
- ~line speed even when 3 subs per node for 13 nodes
- Degradation when 6 pubs and more due to local copies

Experimental Evaluation: Slow Path



100.000 adverts under single scope

- Subscribers subscribe to random item, wait until receive it and reiterate (500 times)

-> worst case for slow path (ignores any possible optimizations due to domain-local rendezvous or mutable semantics)

Node-local

- No net delays
- No TM
- 20ms for 500 processes

Domain-local (Gbit LAN)

- Centralized TM
- ~400 ms for 500 processes per node (7000 subscribers)

Domain-local (PlanetLab)

- Large delays
- ~200ms for 1 sub per node (73 in total)
- ~680ms for 36,500 subs

What is the Take-Away Here?

- In the light of the continuous technological improvement of the Internet, the claim of improving content delivery can be NEITHER a convincing argument for NOR the true potential of information-centric networking!

- Information-centric networking is about utilizing the entire design space provided by storage as well as computation

- To get there, we need to re-think how we design/build systems

PURSUIT has done just that and provided artefacts to demonstrate the benefits

More Information

- Websites
  - http://www.psirp.org (the start of this work)
  - http://www.fp7-pursuit.org (the continuation of this work)
  - http://www.named-data.net/ (successor of CCN)
- Papers
  - ACM CCR 04/2010, SIGCOMM 2009 (LIPSIN), CONEXT 2009 (CCN), and many more on http://www.psirp.org
- Contact: dot@tecvis.co.uk (for questions or student projects)