Video Replication and Access over Fog-based Architecture

Chang, Zhangyu

Supervised by Prof. Gary Chan

Contents

Impact of Video Traffic on Internet

Video contributes to most of the internet traffic

- the weight is continuously increasing
- 73% in 2016 / 82% by 2021 (estimated)

Huge network resource demand for video service

- Overall IP traffic grows 24% percent annually
- Busy-hour Internet traffic is growing more (51% in 2016)

Traditional Approach:

- Content Distribution Networks (CDNs): Not cost-effective
- Peer-to-Peer Network (P2P): Not reliable
- New paradigm to offload the demand

Video Service over Fog

Fog devices:

- E.g., routers, Wi-Fi Aps, set-top boxes, base station
- Lightweight but decent power
- Ubiquitous and close-to-user

Fog-based distribution reduces:

- Load of cloud server
- Server-to-edge traffic
- Inter ISP traffic

Video Popularity Characteristics

Popularity

- Very skewed for both *professionally generated content* (PGC) and *user generated content* (UGC)
- By storing only 10% of long-term popular videos, a cache can serve 80% of requests

Freshness

- The popularity of hot videos decays very quickly
- PGC: 90% of the most popular videos traffic are new each day
- UGC: difficult to predict the popularity of new content

Daily Pattern

- 2 peaks: 2 P.M. and 10 P.M. every day
- lowest at around 5 A.M.: good time to push new content

Challenges for Fog-based Schemes

Distributive

- Fog devices are huge in number and have to collaboratively serve the users

Geography-aware

- It is important to utilize the close-to-user feature to serve the neighbor

Popularity-aware

- Global popular contents have to be pushed into the fog

Lightweight implementation

- Fog device cannot match dedicated server on computation power

Major Problems

Video Replication

- Fog devices are huge in number, but each fog device has only limited storage size
- Ephemeral: highly demanded video for a certain duration and then the demand fades
- What to push and when to push the new contents

Uncoordinated versus *Coordinated* replication schemes

- Uncoordinated: based on device's demand
- Coordinated: based on global popularity

Video Access

- A fog device cannot have full replication of all the videos
- Many fog devices may have the same contents
- Decision affects all users share the same resource

Wired versus *Wireless* fog

- Wired: reduce the problem size
- Wireless: choice of base stations

Solutions

Replication schemes

- Uncoordinated: Variations of LRU / Score-Based Schemes
- Coordinated: Popularity-Based / Division of Storage

Video access schemes

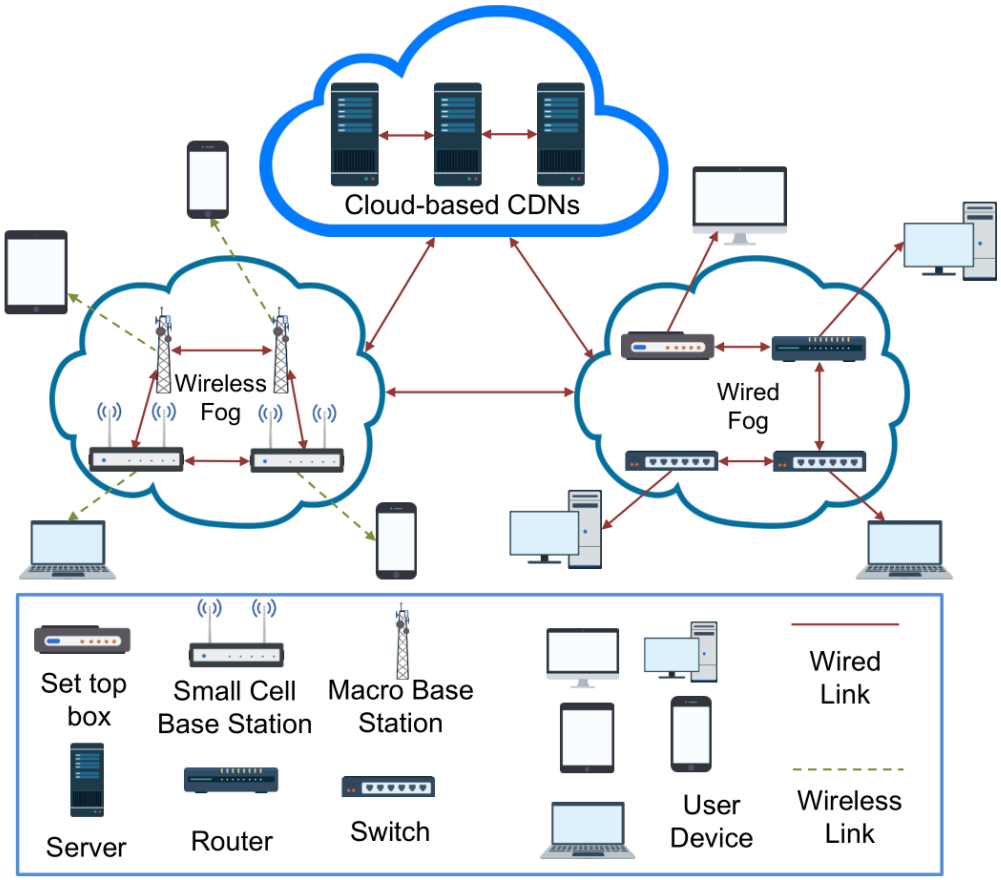
- Wired Users: Clustering Methods / Game Theory Approaches
- Wireless Users: Approximation Algorithm / Mathematical Programming

Case studies

- Youku: CDN Based on Smart Routers
- Thunder Crystal: Crowdsourcing Content Distribution

Contents

System Architecture of Fog-based Video Network



Cloud Layer (Level 1)

- Push new videos to the fog
- Reduced number of servers

Fog Layer (Level 2)

- Backbone of the platform
- Huge in number and close to the users
- Owner and operator of the fog may not be same
- Wired & Wireless fog for all the users

User Layer (Level 3)

- Usually get served by fog devices
- Cloud as the last resort

Comparison: versus CDN and P2P

Table 2.1: Different paradigms for video distribution.

	Cloud CDN	P2P	Fog
Data Storage	Centralized	Distributed	Distributed
System Control	Centralized	Uncoordinated	Coordinated
QoS	Yes	No	Yes
Capital Cost	High	Low	Low
Scalability	Low	High	High
ISP Friendly	Yes	No	Yes
User Contribution	No	Required	Desirable

Comparison: Merits of Fog

vs CDN	vs P2P
Better proximity • Fog device close to the users • Cloud only push to some fog devices	Service guarantees • Full control over fog devices • Stable network connections • Reduce peer churns
Reduced cost • Major cost: real estate, power, cooling and human resource • Do not exist on fog	Coordinated topology • More manageable topology • Reduce the inter-ISP traffic • Better Geography-awareness
User contribution • Users are willing to buy fog devices for better service • Expansion with little cost	Free-riding prevention • Store the contents even the owner is not interested in them • Owner has less turn-off incentive

Contents

Replication Schemes

Replication scheme for fog must be distributive, responsive and easy to implement

Uncoordinated replication schemes

- Each fog device only makes its own decision independently
- Lack global popularity information
- Optimality analysis relies on assumptions

Coordinated replication schemes

- A central server will offer global information to each fog device in some way
- Push the contents directly
- May ignore the local preference

Uncoordinated Schemes

Variations of LRU

- q-LRU: stores the new video content with a probability of *q*
- k-LRU: Storage is divided into k hierarchical part, contents demoted step by step until removal
- Better than LRU if the video popularity distribution follows the Zipf's law
- Assumptions for optimality: Poisson distribution of request & stead state

Score-Based Schemes

- Age-Based Threshold (ABT): calculate a time to live based on access frequency
- Information Centric Network: video has special index, frequency based score but the recent access has higher weight

Uncoordinated Schemes

Scheme	Objective	Parameter to optimize	Methodology	Comment
LRU-based	Maximize hit probability	Video stored in the device	Poisson Approximation	Assume Zipf's Law
iProxy	Improve hit probability	Video stored in the device	Heuristics	New coding scheme
Age-based Threshold	Maximize hit probability	Video lifetime in the device	Poisson Approximation	Assume Zipf's Law

Fog devices may not have enough processing power to handle all the hit

Coordinated Schemes

Popularity-Based Schemes:

- Video content can be proportional to the video popularity
- Deficit bandwidth performs better than proportional replication if fog devices has heterogeneous
- Scheduler gives replication probability *p* or time to live *t* and broadcast this parameter to all the fog devices
- New indices (geographic/propagation/social influence)

Division of Storage:

- Divide the fog storage into 2 parts
- Local/global popularity & PGC/UGC

Coordinated Schemes

Scheme	Objective	Parameter to optimize	Methodology	Comment
Global Popularity	Reduce server load	Video stored in the device	Mathematical modelling	Not consider bitrate/device capacity
Deficit Bandwidth	Reduce server load	Video stored in the device	Mathematical modelling	Consider device capacity
Last-mile Implementation	Reduce traffic cost	Probability to store a video	Primal-dual approach	Comprehensive model
Social Video Index	Improve hit probability	Video indices	Heuristics	Based on measurement
Division of Storage 1	Maximize hit probability	Storage Division	Heuristics	Support Wired & Wireless users
Division of Storage 2	Maximize social welfare	Storage Division	Supermodular game	Support social video

Contents

Video Access Schemes

Video access decision is critical to effectively utilize fog network resources and avoid network congestion

- Multiple replications for the same video content in several fog devices

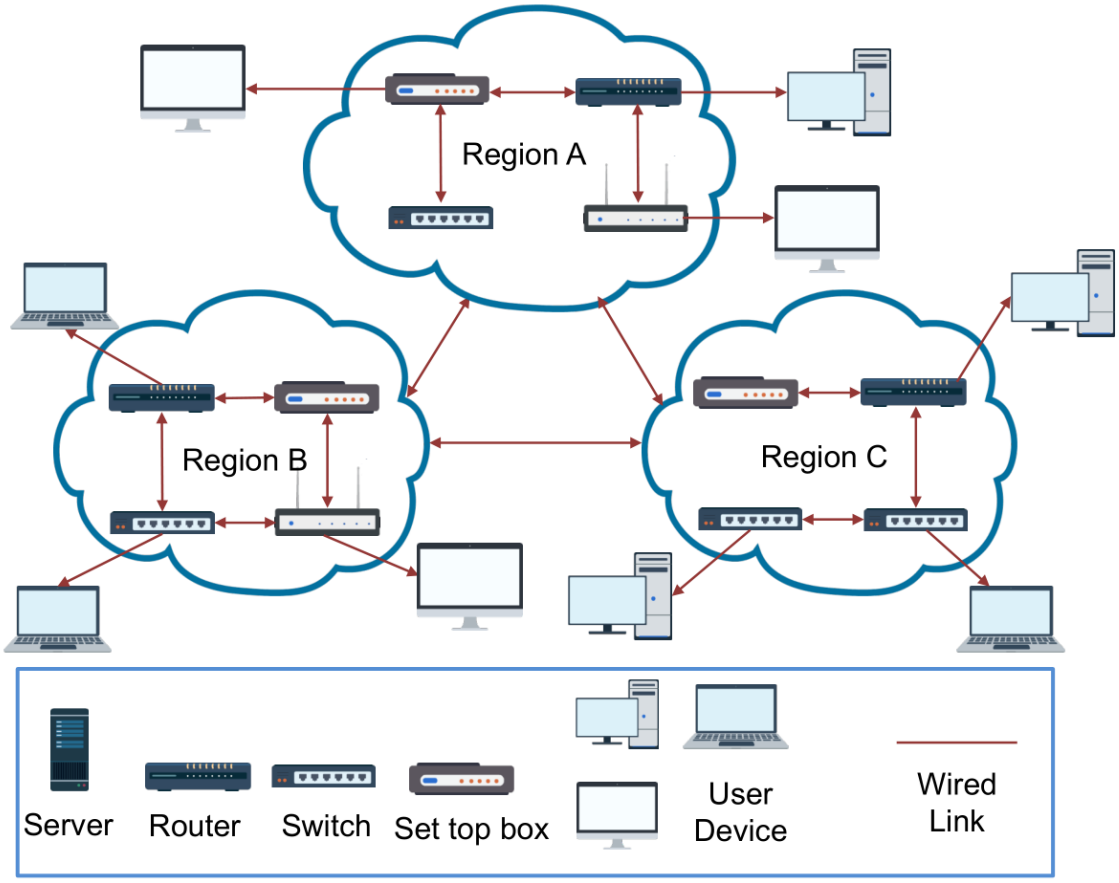
Huge size of the problem to optimize the whole network

- Number of both the videos and fog devices are huge
- Facility Location Problem: NP-complete

Effectively Solution

- Reduce the problem size (clustering/ divide and conquer)
- Find the approximation algorithms

Wired User



Thousands of devices

- NP-hard in nature
- Divide and conquer
- Brutal force for small-scale

Clustering Methods

- Geographic location/ISP
- Similarity of popularity
- Auction-based method with each region
- MCMF problem for inter-region traffic

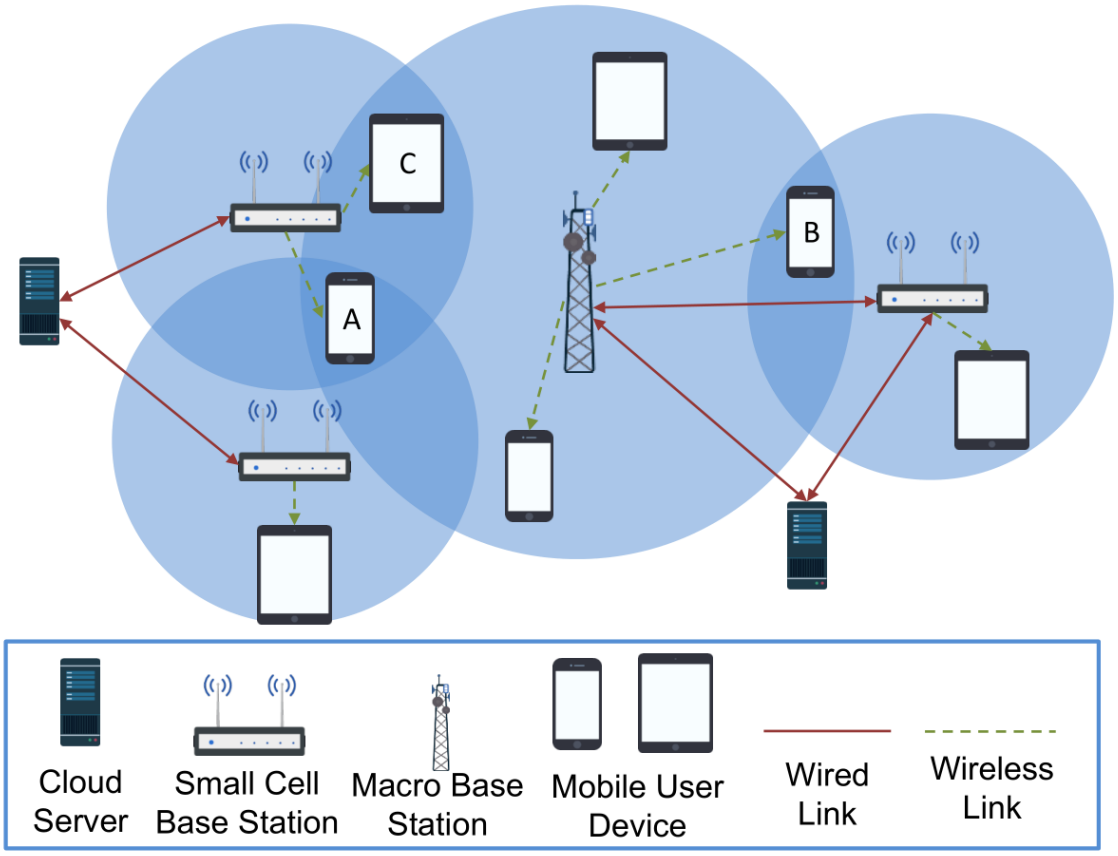
Game Theory Approaches

- Between device owners & operators (Stackelberg Game)
- Between social friends (supermodular game)

Wired User

Scheme	Objective	Parameter to optimize	Methodology	Comment
Cluster fog devices 1	Reduce server load	How to partition fog devices	Sampling & greedy algorithm	Based on measurement
Cluster fog devices 2	Reduce server load	Traffic between clusters	Linear programing & Heuristics	2 schemes can be combined
Game theory approach 1	Total revenue	Price to use a fog device	Stackelberg Game	Fog owner and CP can cooperate
Game theory approach 2	Maximize social welfare	From which friend to get video	Supermodular game	Assume friends share videos

Wireless Users



Joint optimization of the replication and access problem

- User may covered by multiple base stations
- Relative small problem size

Approximation Algorithm

- Facility location problem
- Cluster user demand
- Randomized method if user preference differs not much

Mathematical Programming

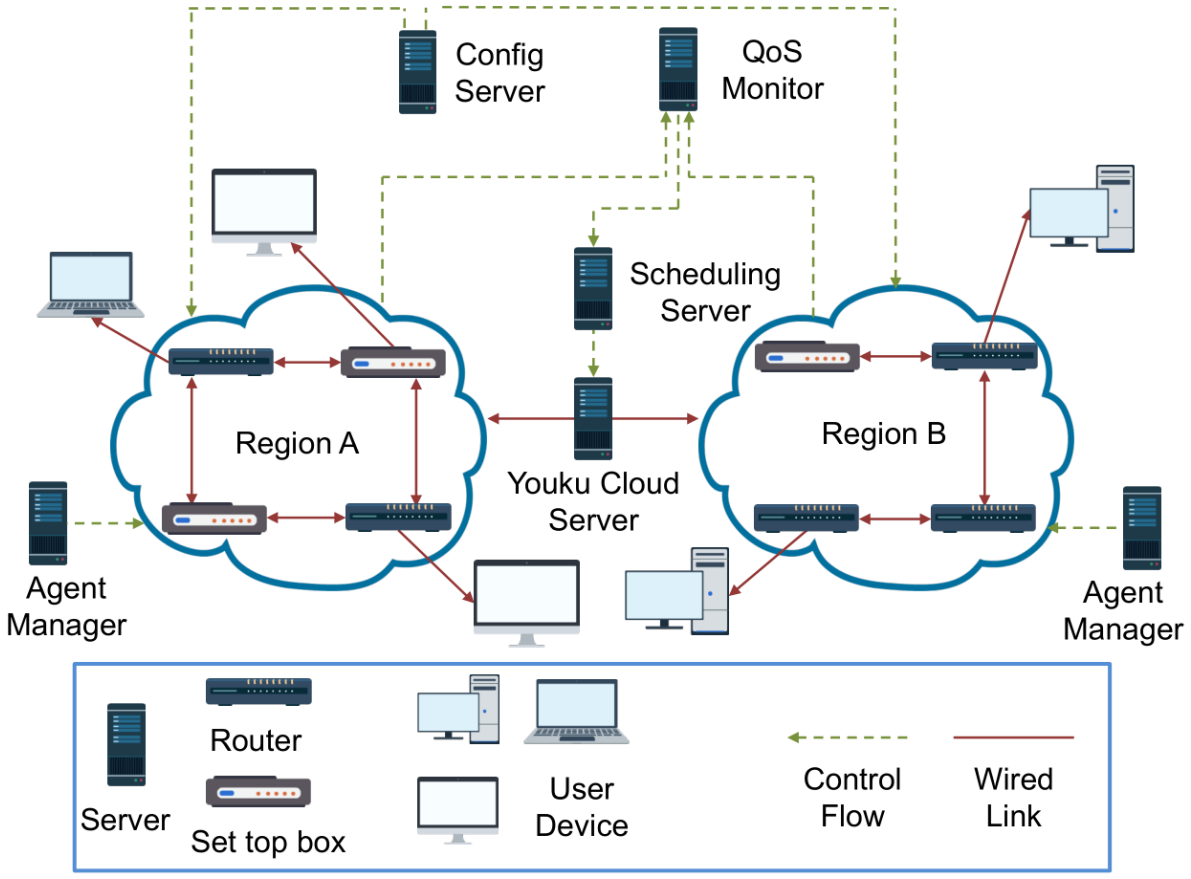
- Allowing use has probability to get the video from different base stations
- Use network coding to bypass the NP-hardness of integer linear programming

Wireless Users

Scheme	Objective	Parameter to optimize	Methodology	Comment
JRC-UR	Minimize server load	Replication & Access	Approximation of LBS	Approximation ratio given
BS assisted D2D	Minimize server load	Replication & Access	Monte Carlo optimization	Heuristics in nature
AP deployment	Minimize server load	Fog device deployment	Integer linear programming	With a greedy heuristics
An online algorithm	Minimize server load	Replication & Access	Convex programming	Allow user to access many APs
FemtoCaching	Minimize server load	Replication & Access	Linear Programming	Use coded content

Contents

Youku: CDN Based on Smart Routers



Peer Video CDN based on Smart-routers :

- Subsidized smart-routers

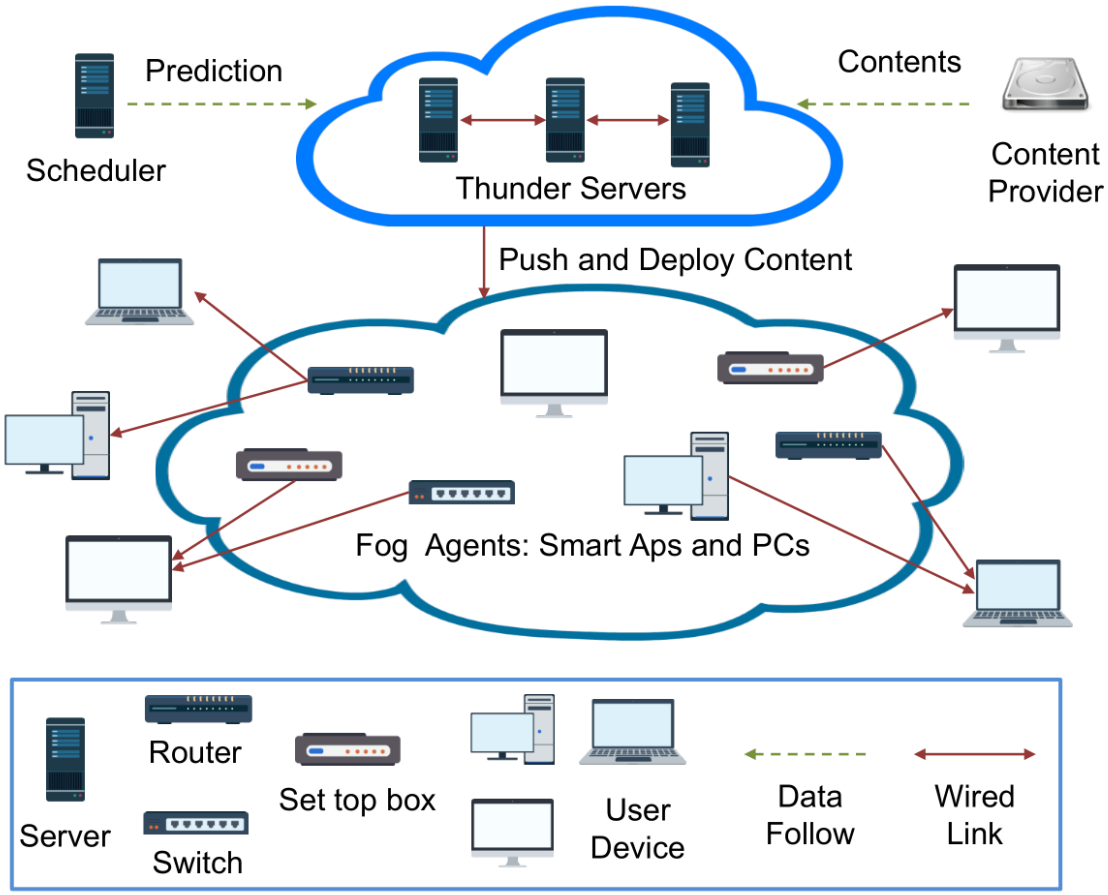
Centralized coordinating

- 4 kinds of servers to manage the smart routers

Combining caching with recommendation

- Push and store home page contents (73%)
- Push between 0 and 3 A.M. (daily pattern)
- Do not care about local popularity

Thunder Crystal: Crowd-sourcing Content Distribution



Crowdsourcing Content Distribution

- Crowdsourcing system
- Device owner get monetary return

Central Managed Pushing

- Push new content for content provider
- Fixed cloud serve budget
- Popularity decays exponentially
- Indiscriminate policy

Random user access scheme

- No geography awareness

Contents

Conclusion

Fog-based content distribution:

- Low operation cost and better quality-of-service

Replication:

- Uncoordinated: simple but no global popularity
- Coordinated: popularity aware but no preference

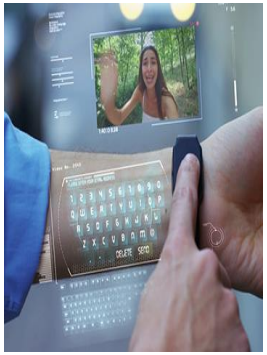
Video Access

- Wired user: divide users into areas
- Wireless user: jointly optimize replication & access

Case Studies

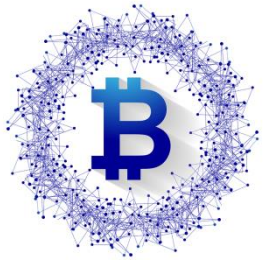
- Youku: combining pushing with recommendation
- Thunder: crowdsourcing with random pushing and access

Future Directions



Augmented Reality
- Combination of real-world and video
- Location-based service
- Fog devices for localization and video distribution

Video Data Analytics
- Video from monitoring camera keeps increasing (7 times from 2016 to 2021)
- Fog devices at the camera side can analyze the video and upload the results or features for less traffic



Blockchain
- Manage the transactions over the Internet
- Negotiation distributively on network control
- Any databases can be saved distributively in the fog

Thank You!

Any Questions?

Comments

Survey title: VoD-Fog Approaches and data driven comparisons

Zhang Qian: Future work shall be consistent with the main problem

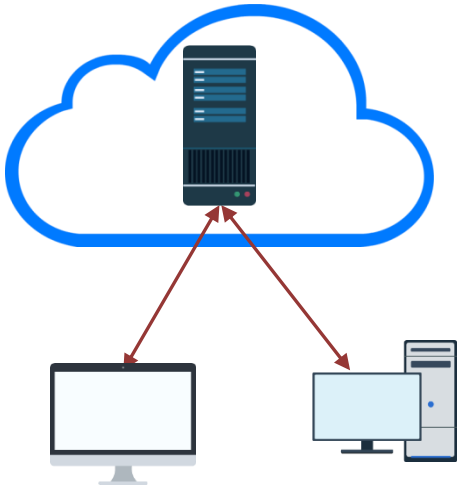
BB: How to differentiate between Fog & P2P? Many approaches have been used in P2P

Many devices (e.g., Set top box) in previous P2P can be regarded as Fog (End user heterogenity)

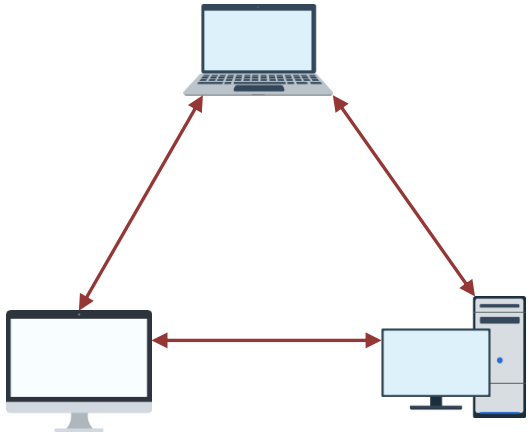
Fog: Large number of medium range devices & Reliable mid-layer for better tractability and billing

Presentation skills : needs more Passion

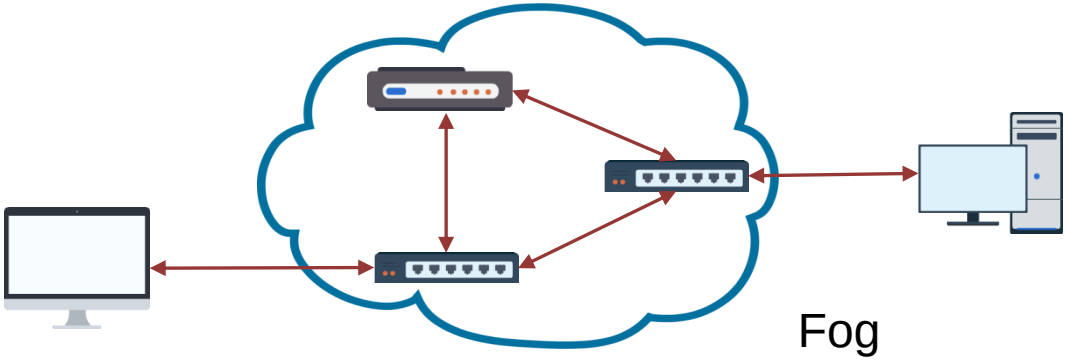
Comments (Cont'd)



Cloud



P2P



Fog