边缘缓存

刘 芳

liufang25@mail.sysu.edu.cn

中山大学 数据科学与计算机学院

目录

背景

- 1998年，Akamai公司，内容分发网络（CDN）
- CDN将用户的访问指向距离最近的缓存服务器上，以此降低网络拥塞，提高用户访问响应速度和命中率。
- CDN强调内容（数据）的备份和缓存，而边缘计算的基本思想则是功能缓存[1]。

百度的CDN节点

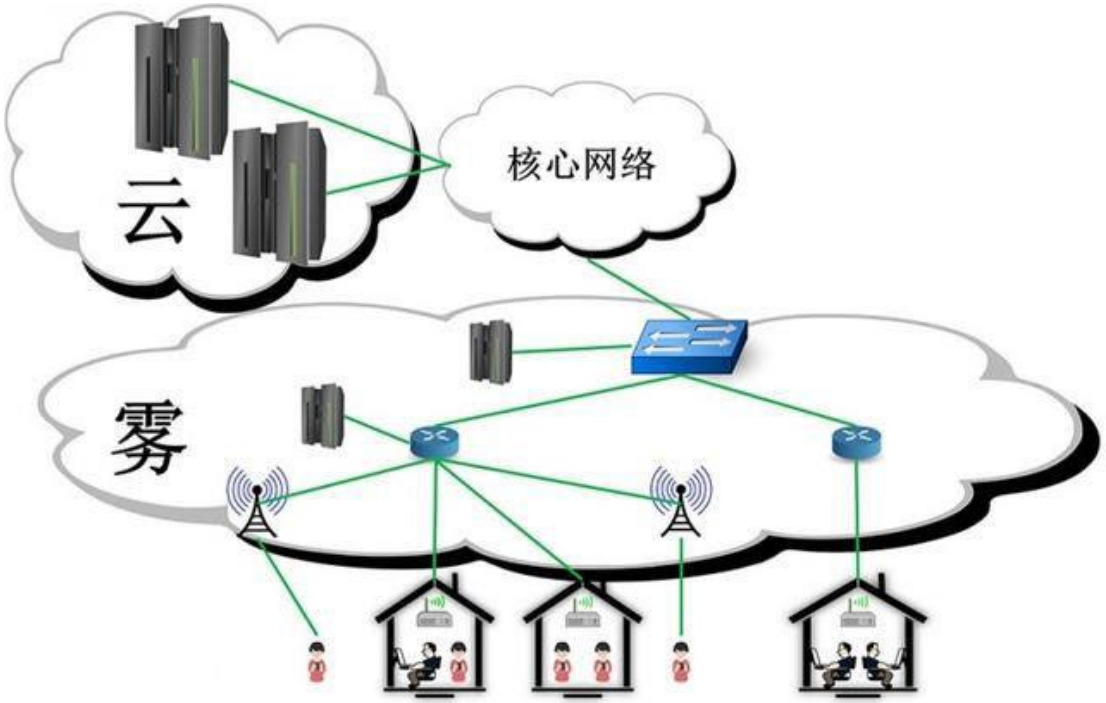
[1]施巍松,张星洲,王一帆,张庆阳. 边缘计算：现状与展望[J]. 计算机研究与发展, 2019, 56(1): 69-89.

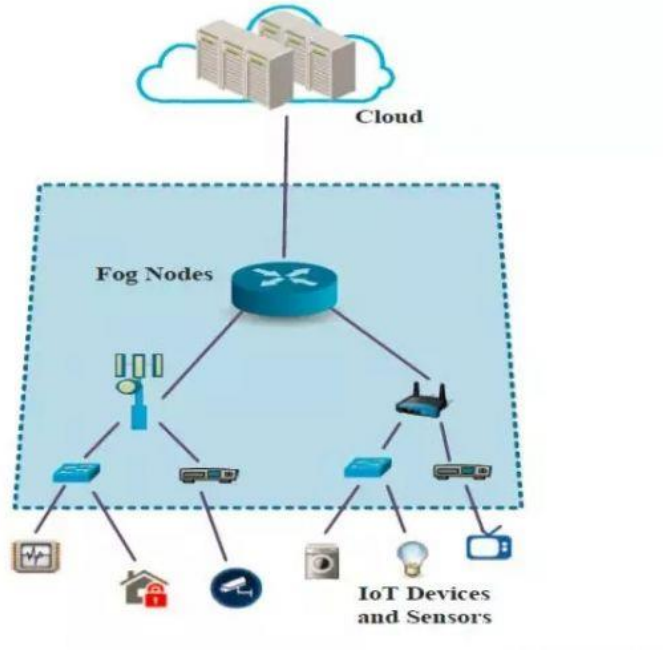
背景

- 2006年，Google，云计算
- 然而这种集中式的云架构对时间敏感、带宽稀缺的工业物联网就不再适用。
- 因此某些关键数据的处理任务最好是在数据源而不是云端，边缘缓存大有可为。

背景

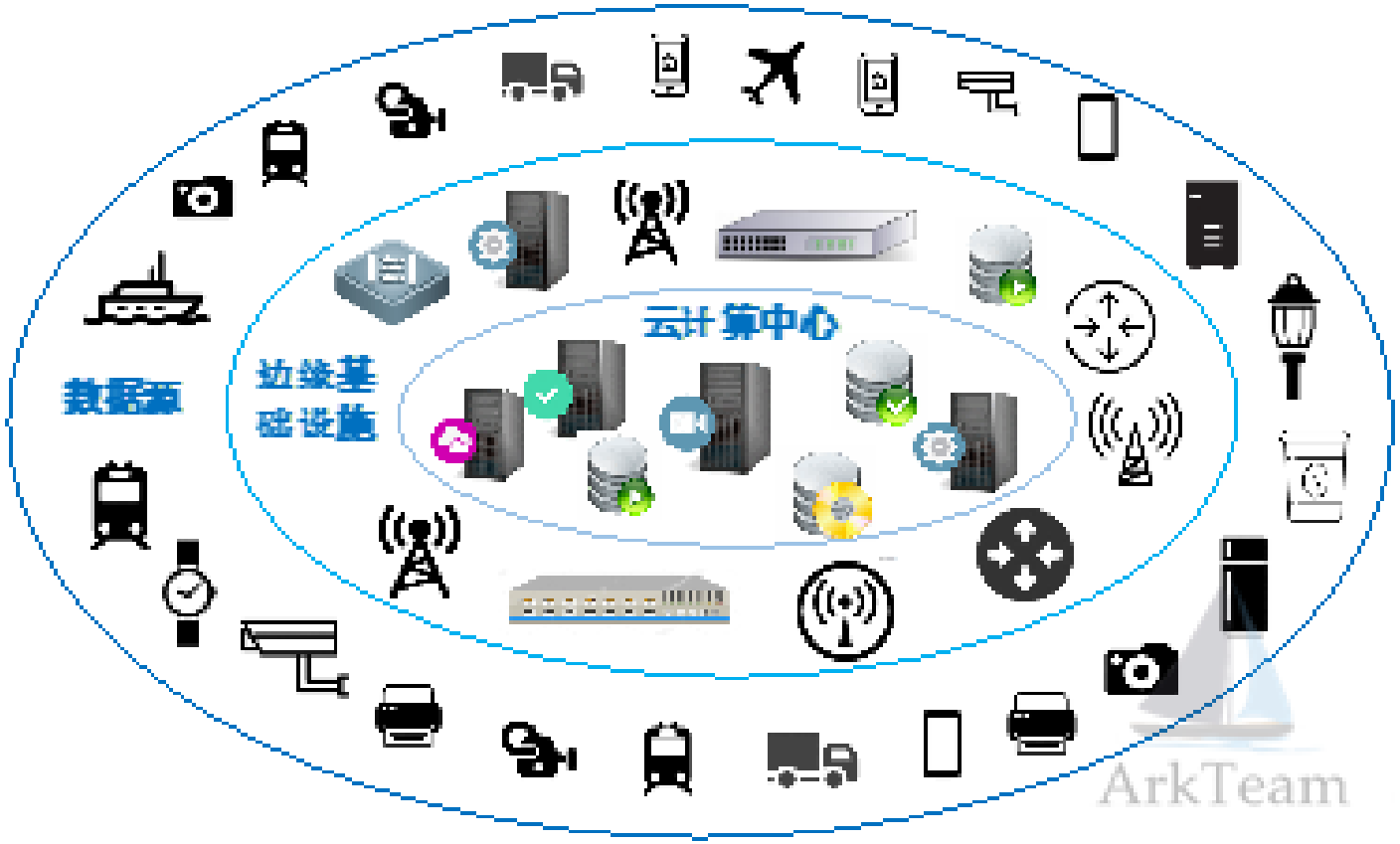
- 2012年，思科，雾计算：迁移云计算中心任务到网络边缘设备执行的一种高度虚拟化计算平台。
- 雾计算关注基础设施之间的分布式资源共享问题
- 而边缘计算除了关注基础设施之外，也关注边缘设备，包括计算、网络和存储资源的管理，以及边端、边边和边云之间的合作。





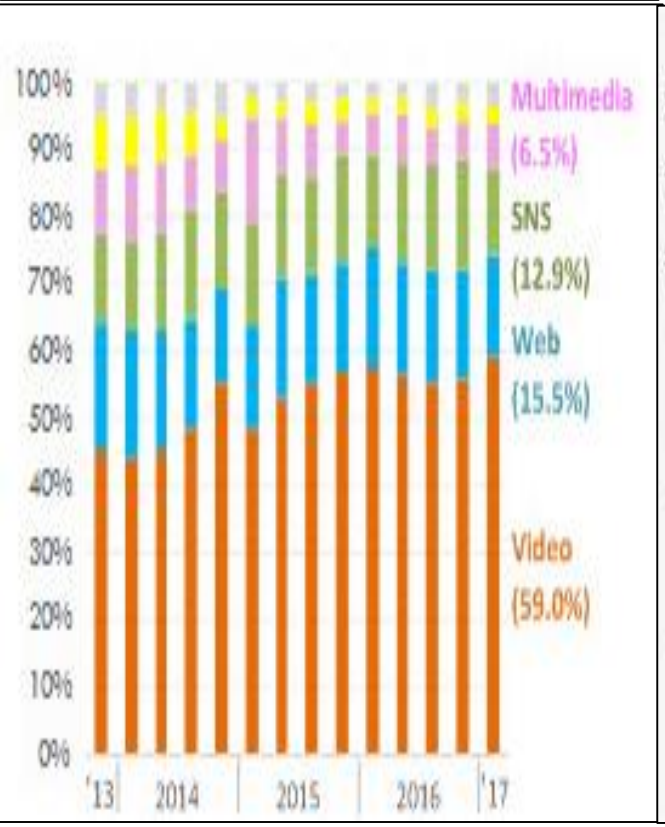
背景

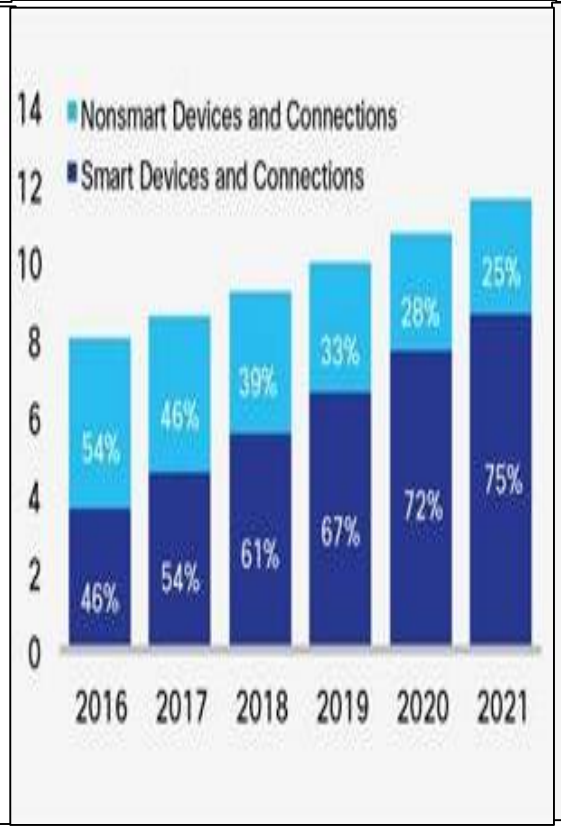
- 2013年，美国太平洋西北国家实验室，边缘计算。
  - 承载云服务功能的下行，万物互联服务的上行。

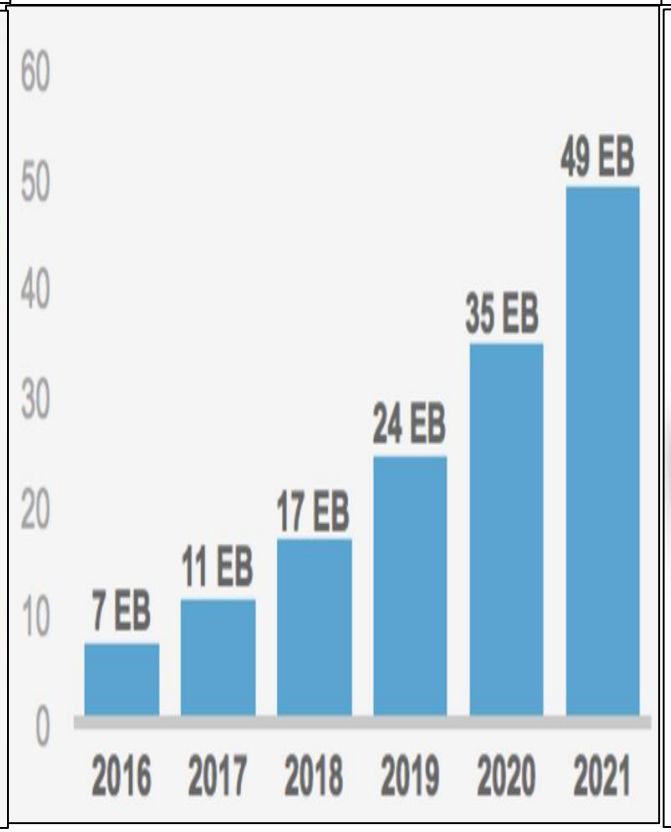


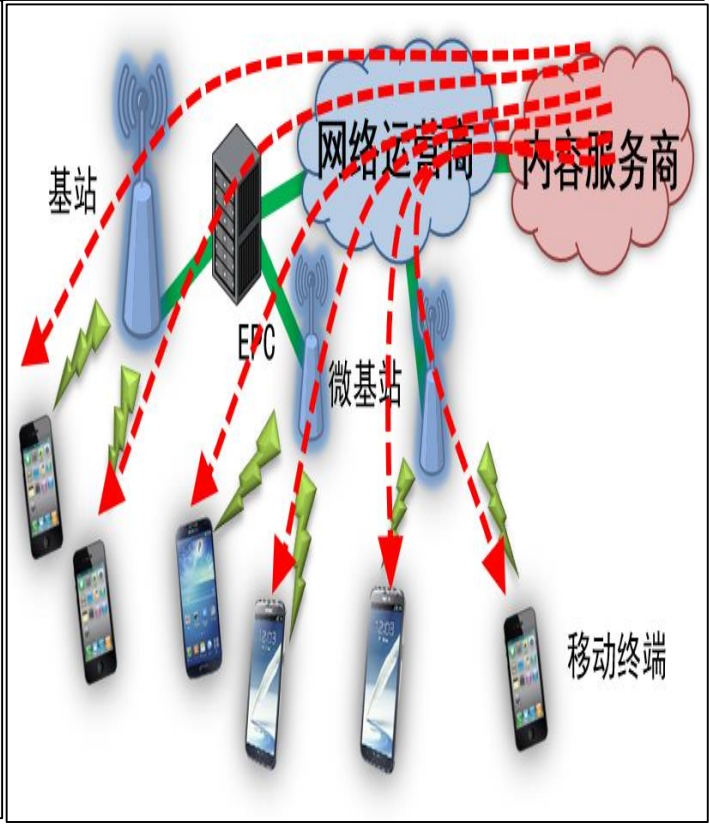
背景

高质量内容爆增	设备数猛涨	移动流量爆炸	重传比例极高





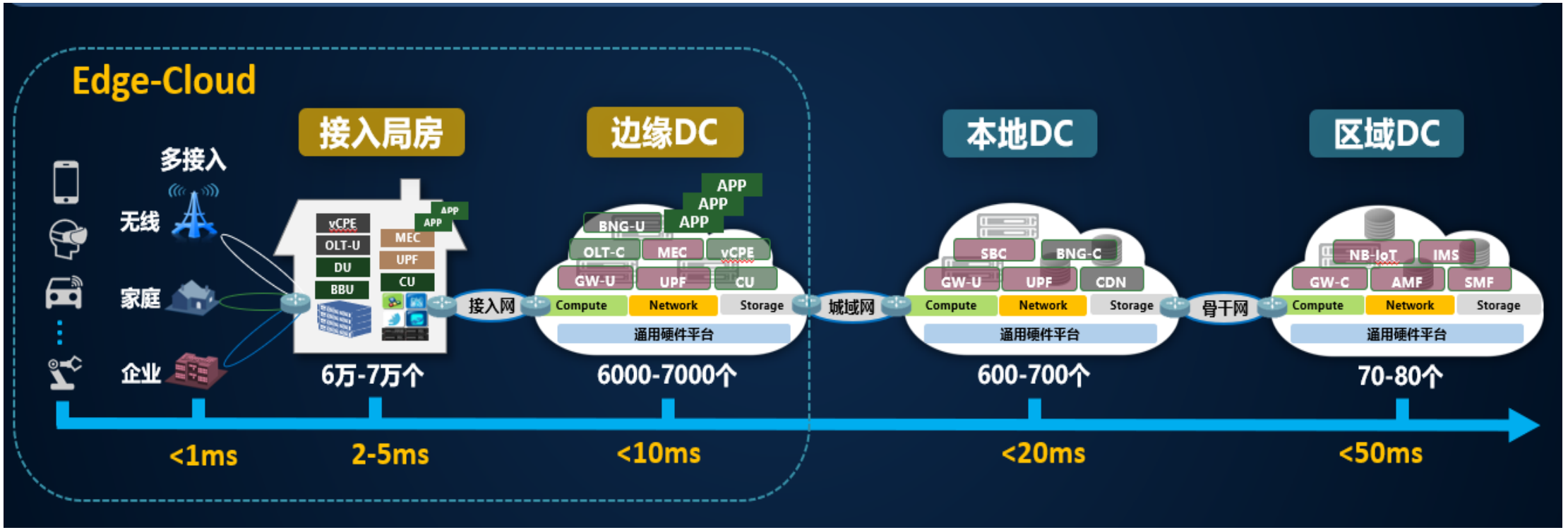




[CISCO VNI移网络流量预测’16][Softbank软银移动网络流量报告’17][Shinae‘14][Claudio’16]等论文和报告

背景

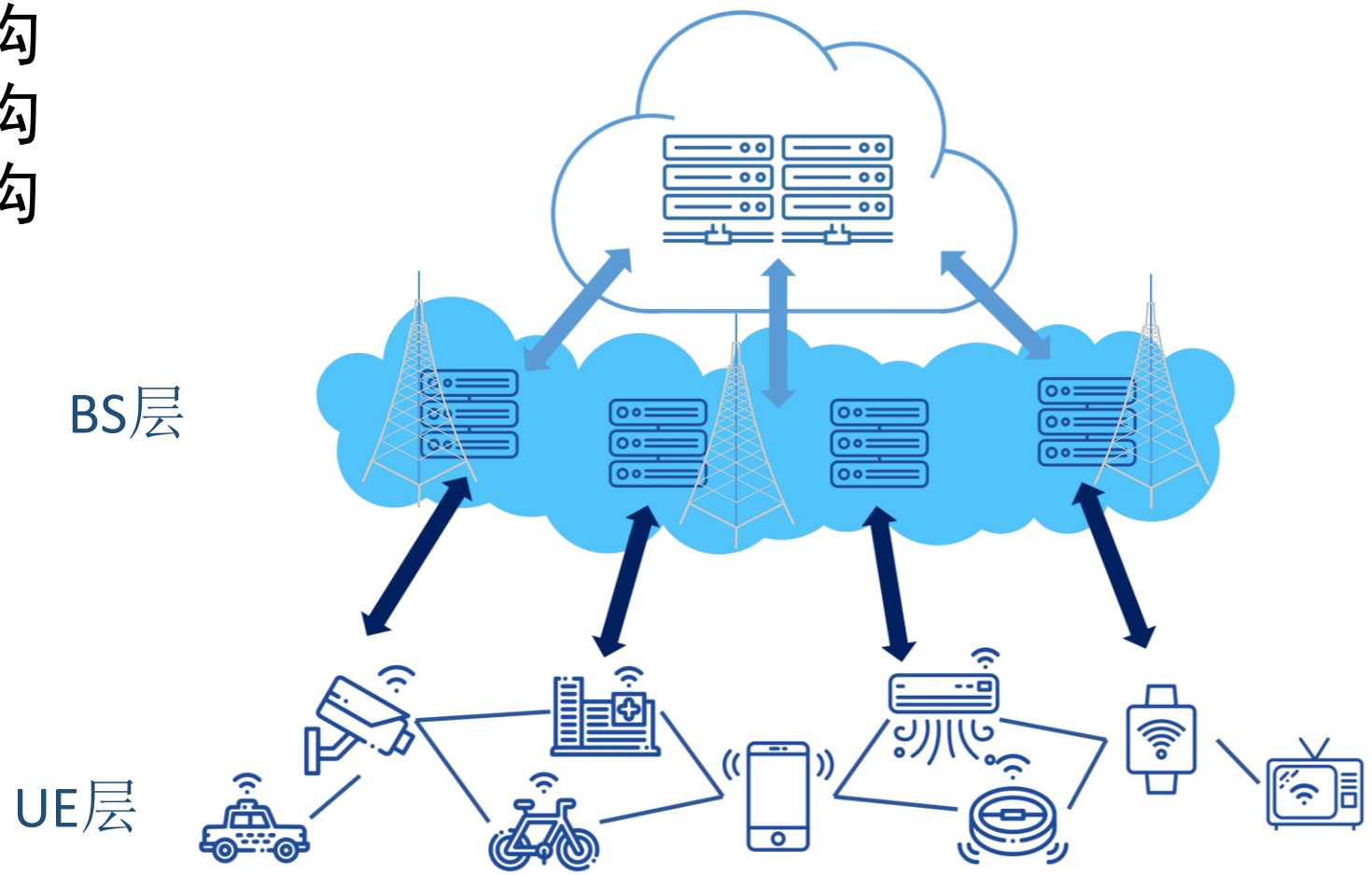
- 据IDC的预测，到2019年，IoT产生的数据将有45%需要在网络边缘进行存储、处理、分析和操作。
- 到2020年将有超过500亿的终端与设备联网，未来超过50%的数据需要在网络边缘侧分析、处理与储存，故边缘缓存所面对的市场规模非常巨大。
- 边缘缓存是降低无线接入网(RANs)中繁重的通信量和端到端延迟的一种有效方法。



目录

分布式架构

- 边缘缓存的分布式架构，根据缓存内容的放置，可分为[1]：
  - BS层缓存的分布式架构
  - UE层缓存的分布式架构
  - 多层缓存的分布式架构



[1] Zhuying Piao, et al. Recent Advances of Edge Cache in Radio Access Networks for Internet of Things: Techniques, Performances, and Challenges. IEEE INTERNET OF THINGS JOURNAL, VOL. 6, NO. 1, FEBRUARY 2019

分布式架构-BS层

- Request -> helper? -> MBS? -> CN

- 若helper之间的距离较大，一个移动设备只连接到一个helper，则helper根据所请求内容在连接的移动设备之间的流行度分布缓存内容。
- 若helper密集部署，一个设备可以连接到多个helper。则这些helper分别缓存该用户设备请求的流行内容，以最小化平均延迟。

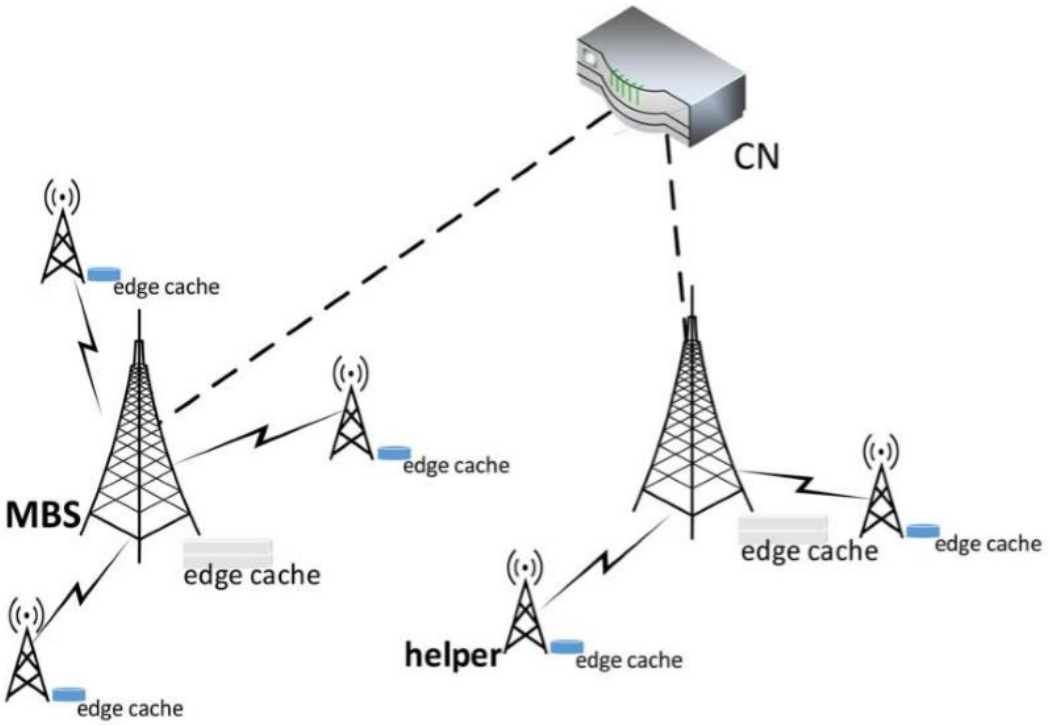


Fig. 2. Edge cache in HetNet.

V. Chandrasekhar, J. G. Andrews, and A. Gatherer, “Femtocell networks: A survey,” IEEE Trans. Wireless Commun., vol. 46, no. 9, pp. 59–67, Sep. 2008.

分布式架构-BS层

- 每个设备都可以由一个或多个SBSs提供服务，这取决于它的位置和BS密度。
- 如果设备只由一个SBS提供服务，则SBS将直接将缓存的请求内容传输给该设备。否则，SBS将向CN发送请求以获取内容。
- 如果设备由多个SBS提供服务，则可以使用协同波束赋形将内容从不同的SBSs传输到设备。

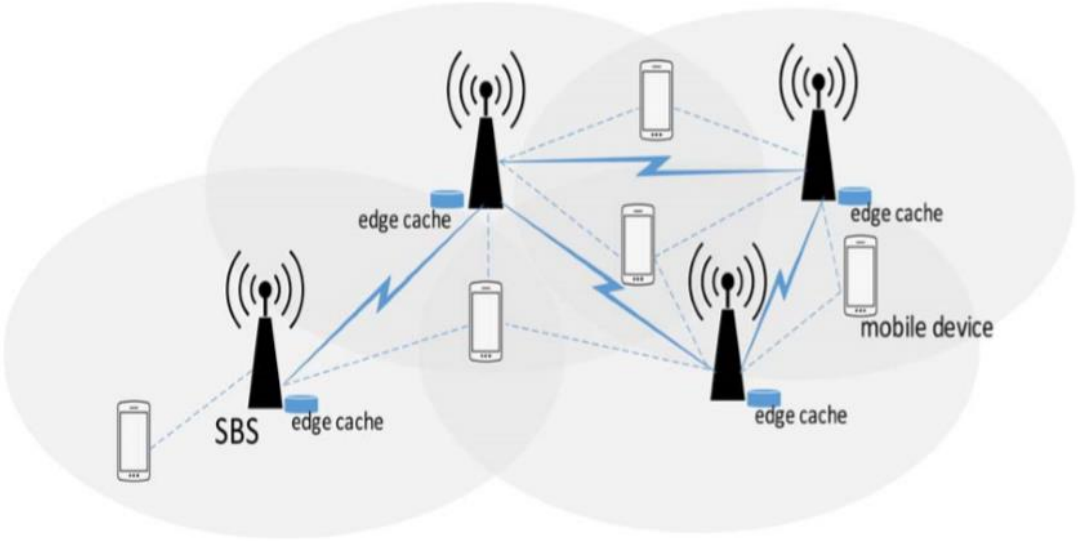


Fig. 3. Edge cache in self-organized RANs.

J. Liu, B. Bai, J. Zhang, and K. B. Letaief, “Content caching at the wireless network edge: A distributed algorithm via belief propagation,” in Proc. ICC, Kuala Lumpur, Malaysia, Jul. 2016, pp. 1–6.

分布式架构-UE层

- 设备与设备，D2D通信，线性拓扑；

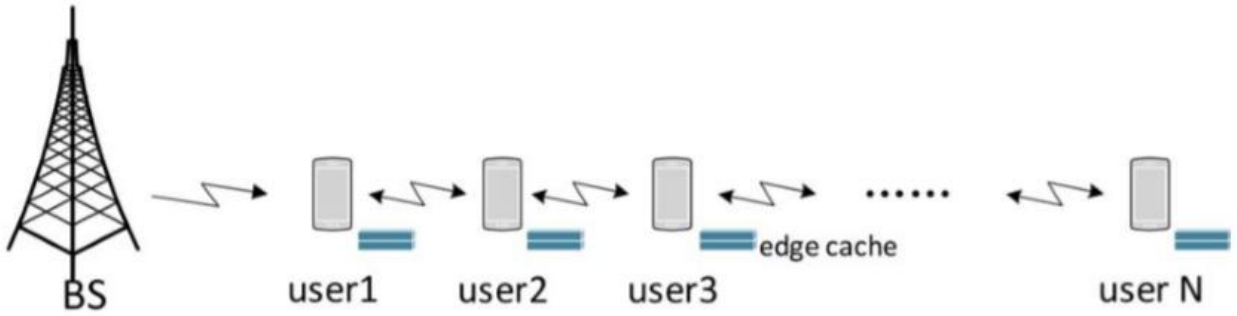


Fig. 4. Wireless multihop linear topology.

- 社交距离描述的是在一个简单的社交空间中，用户对内容的兴趣程度或两个用户之间的亲密程度。距离越小，兴趣或关系越密切。

- 例如，用户3比用户2对内容3更感兴趣，用户2和3之间的关系比用户5和6之间的关系更密切。

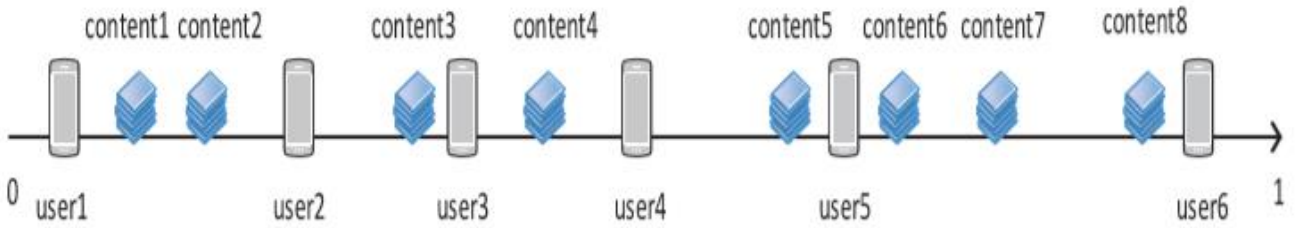


Fig. 5. Social space with users and contents.

J. Cho, S. Oh, J. Kim, H. H. Lee, and J. Lee, “Neighbor caching in multi-hop wireless ad hoc networks,” IEEE Commun. Lett., vol. 7, no. 11, pp. 525–527, Nov. 2003.

分布式架构-多层

- BBUs: Baseband Units
- RRHs: Remote Radio Heads
- 带雾计算和边缘缓存功能的RRH称为边缘计算接入点(ECAP)

- F-RAN中的四种传输模式：
  - D2D Mode
  - Nearest EC-AP Mode
  - Local distributed coordination Mode
  - Global C-RAN Mode

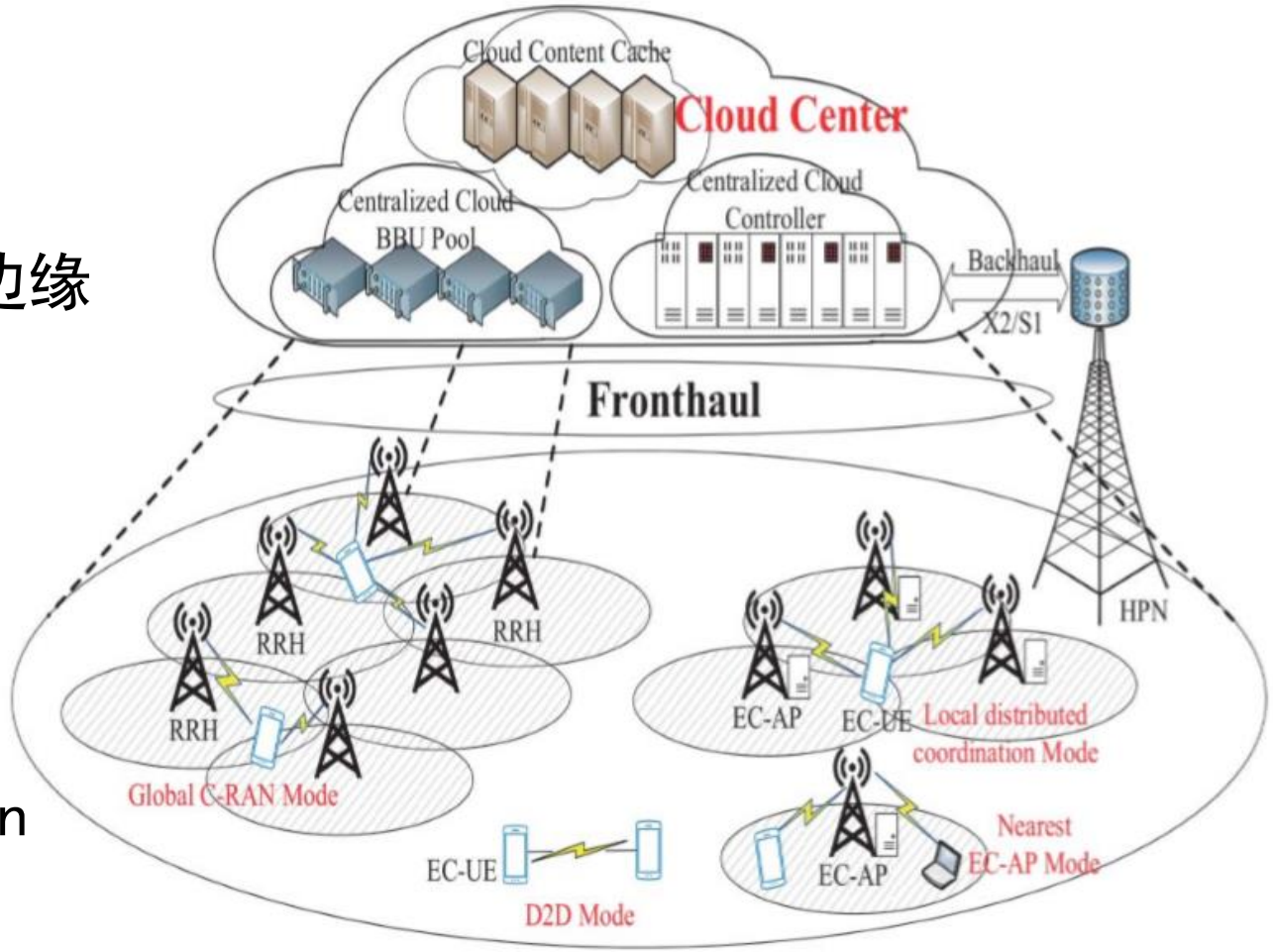


Fig. 6. System architecture of the F-RAN [47].

M. Peng and K. Zhang, "Recent advances in fog radio access networks: Performance analysis and radio resource allocation," IEEE Access, vol. 4, pp. 5003–5009, 2016.

分布式架构-多层

- 内容放置阶段
  - CU分析DUs流行信息，并主动缓存
  - DU根据其范围内的流行信息从CU获取内容进行缓存
- 服务请求阶段
  - Request -> DU? -> CU? -> CN

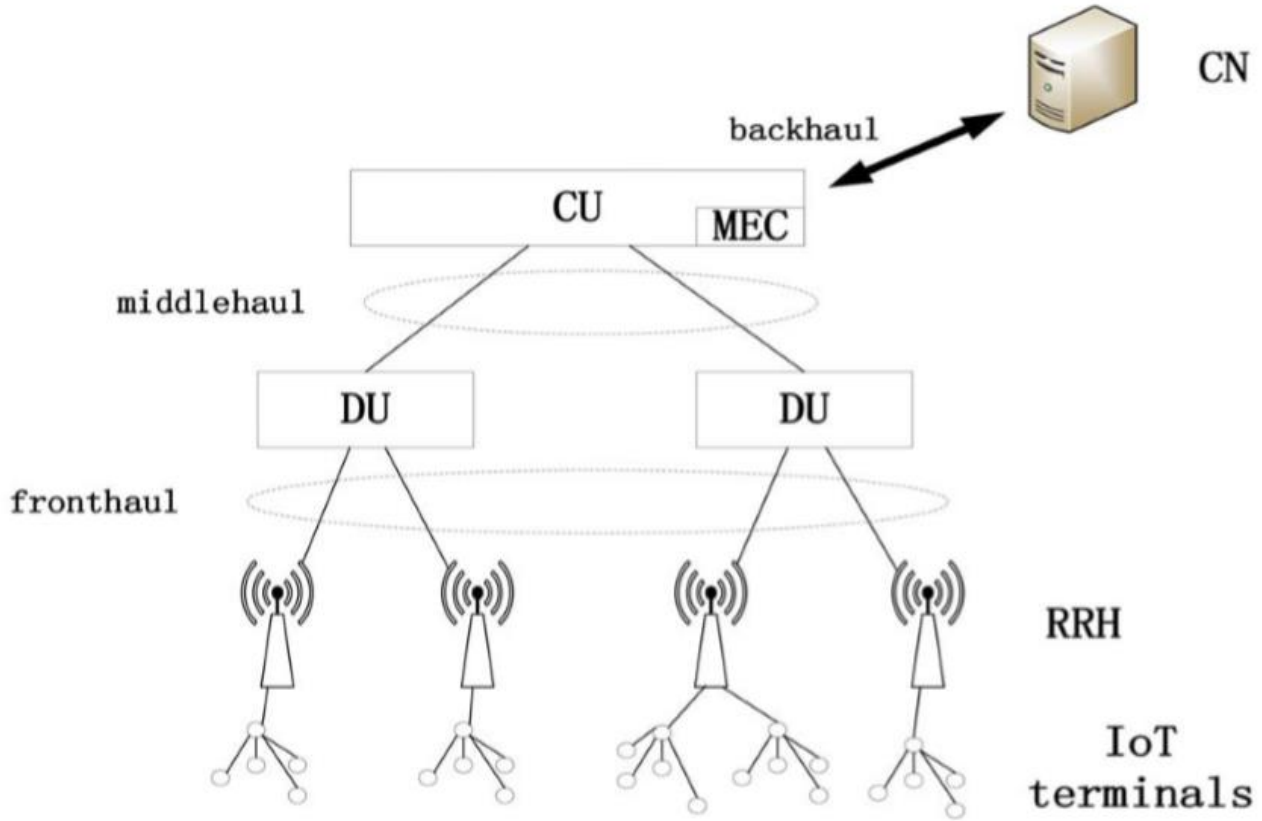


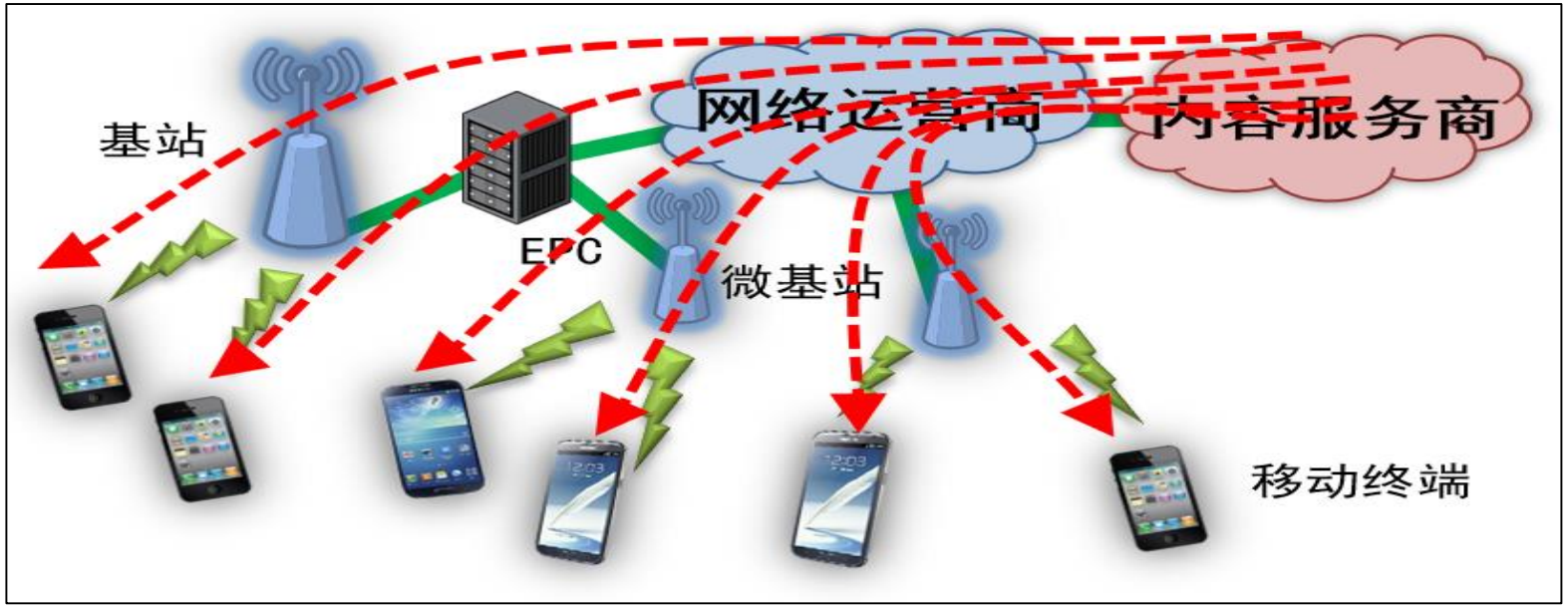
Fig. 7. Hierarchical edge cache structure.

Zhuying Piao, et al. Recent Advances of Edge Cache in Radio Access Networks for Internet of Things: Techniques, Performances, and Challenges. IEEE INTERNET OF THINGS JOURNAL, VOL. 6, NO. 1, FEBRUARY 2019.

目录

协同机制

- 随着网络用户规模的不断增长，传统的服务器－客户端或者服务器－蜂窝网基站－客户端的传输架构会产生服务器负载过重、传输链路上的单点失效、以及关键链路上热门视频的重复冗余传输等问题。
- 服务提供商将面临巨额的带宽开销，同时用户的服务质量也难以保证。
- 目前存储设备的容量和成本都在大幅度下降，缓存容量相对于庞大的用户需求仍然相对不足，因此单个基站和用户的缓存往往无法满足实际需求。
- 在提升单个缓存利用率的基础上，如何利用多个用户和多个基站协作缓存，也成为缓存研究中的重要问题。



协同机制

- Ammar Gharaibeh等人考虑了小型基站之间的协同缓存机制，并设计了一种高效的在线协作缓存部署与替换方法[1]。

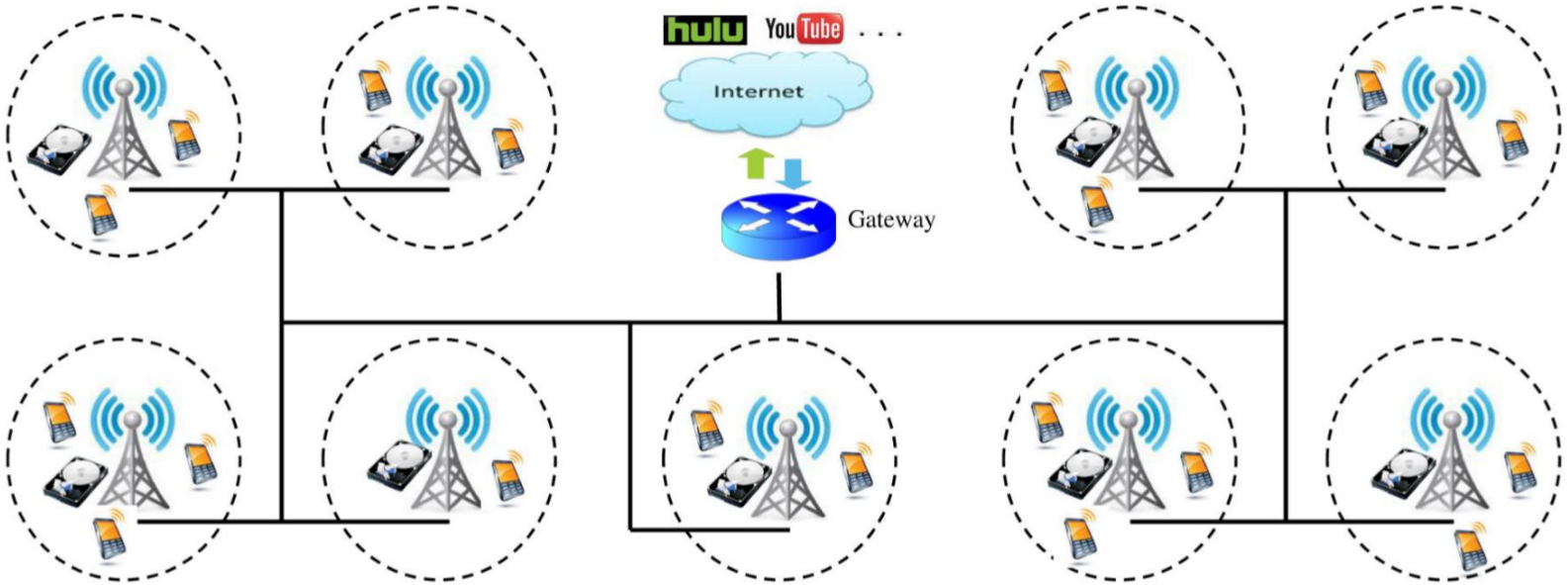


Fig. 1: Collaborative multicell-coordinated system.

[1] Gharaibeh A, Khreishah A, Ji B, et al. A provably efficient online collaborative caching algorithm for multicell-coordinated systems[J]. IEEE Transactions on Mobile Computing, 2015, 15(8): 1863-1876.

协同机制

$$\min \sum_{i=1}^{K}\sum_{k=1}^{K+1}\sum_{j=1}^{M} T_{ij}^{k} X_{ij}^{k} \gamma_{ij} s_j + \sum_{i=1}^{K}\sum_{j=1}^{M} f_{kj} Y_{kj} s_j$$

Subject to

$$X_{ij}^{k} \le Y_{kj}, \quad \forall i,j,k \tag{1}$$

$$\sum_{k=1}^{K+1} X_{ij}^{k} \ge 1_{\{\gamma_{ij}>0\}} \quad \forall i,j \tag{2}$$

- f_{kj}表示内容j在基站k上的单位存储成本；
- Y_{kj}表示内容j是否缓存在基站k中，是则1，否则0；

- T_{ij}^{k}表示基站i从基站k取内容j的用户流失成本，用最短路径进行表征；
- X_{ij}^{k}表示基站i是否从基站k取内容j，是则1，否则0；
- γ_{ij}表示基站i范围内，内容j的请求次数；
- s_j表示内容j的大小。

[1] Gharaibeh A, Khreishah A, Ji B, et al. A provably efficient online collaborative caching algorithm for multicell-coordinated systems[J]. IEEE Transactions on Mobile Computing, 2015, 15(8): 1863-1876.

协同机制

- Li考虑网络设备提供商与视频服务提供商之间的竞争关系，通过斯坦伯格博弈模型（Stackelberg Game）刻画缓存设备租赁及部署策略，且同时最大化网络设备提供商和视频服务提供商的收益。

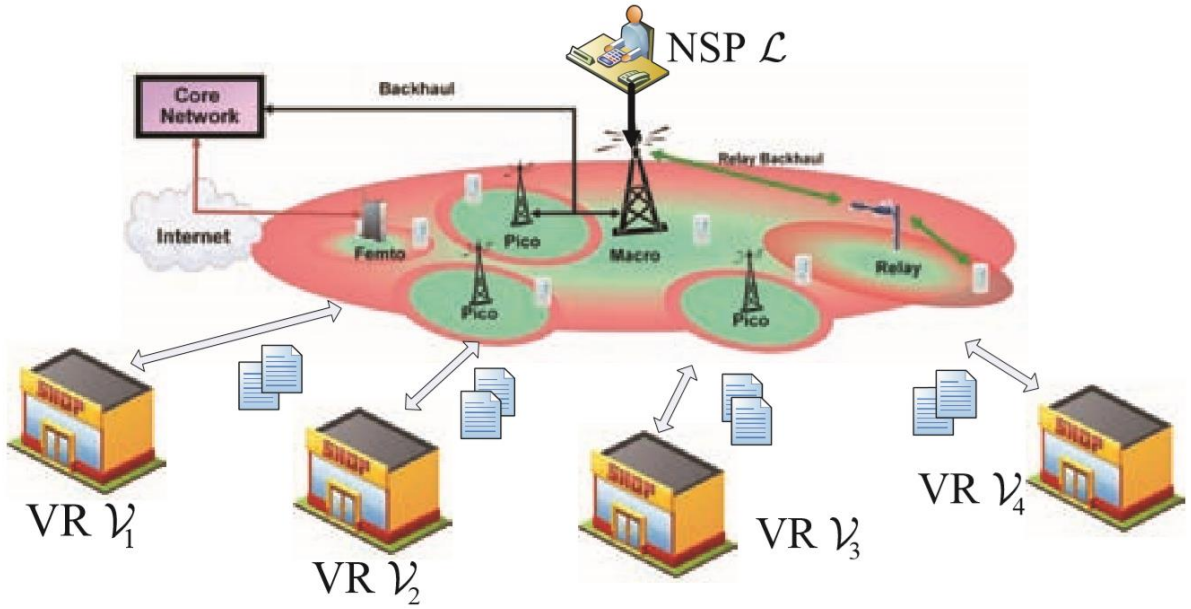


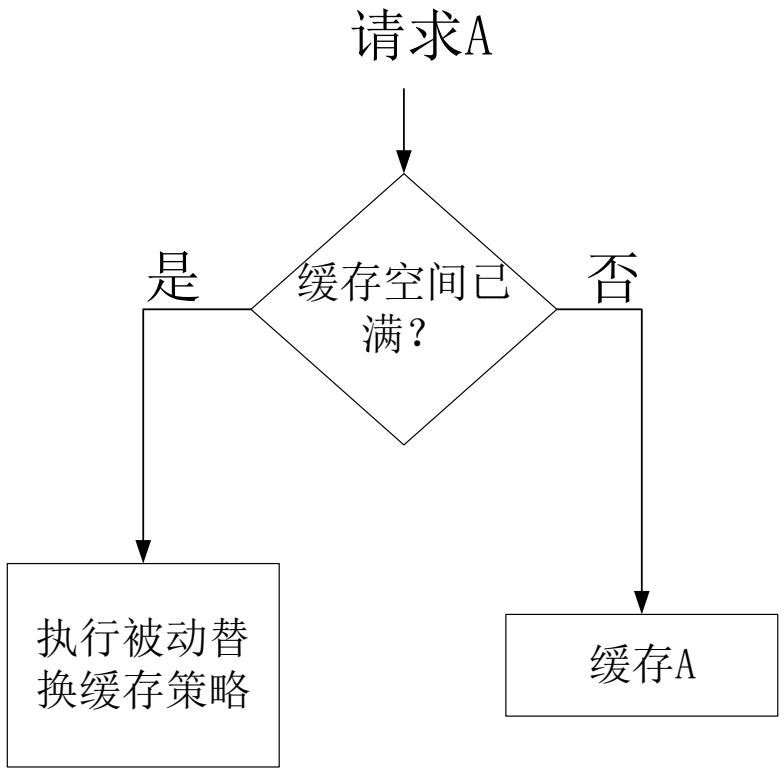
Fig. 1. An example of the small-cell caching system with four VRs.

[3] Li J, Chen H, Chen Y, et al. Pricing and resource allocation via game theory for a small-cell video caching system[J]. IEEE Journal on Selected Areas in Communications, 2016, 34(8): 2115-2129.

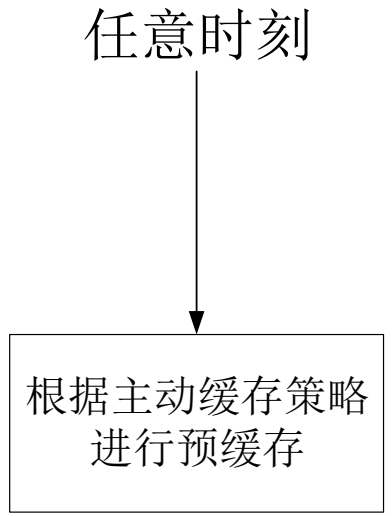
目录

替换策略

- 被动缓存策略
- 主动缓存策略



被动缓存



主动缓存

Zhuying Piao, Mugen Peng , Yaqiong Liu, Mahmoud Daneshmand. Recent Advances of Edge Cache in Radio Access Networks for Internet of Things: Techniques, Performances, and Challenges. IEEE INTERNET OF THINGS JOURNAL, VOL. 6, NO. 1, FEBRUARY 2019.

替换策略-被动

- LRU、LFU、FIFO
- 混合LFU和FIFO[1]：
  - 将缓存空间分成两部分，各自执行LFU和FIFO，分别缓存了最频繁访问和最近访问的内容。

- Neighbor Caching-LRU[2]：
  - 在自组网中，当本地缓存已满，将替换出的内容缓存在邻近（一跳）节点中。
  - 替换策略：根据需要替换出去的内容RUT（最近使用时间），匹配具有最早的RUT的邻近节点，将内容放置在该节点上。

[1] Y. Zhou, L. Chen, C. Yang, and D. M. Chiu, “Video popularity dynamics and its implication for replication,” IEEE Trans. Multimedia, vol. 17, no. 8, pp. 1273–1285, Jun. 2015.
[2] J. Cho, S. Oh, J. Kim, H. H. Lee, and J. Lee, “Neighbor caching in multi-hop wireless ad hoc networks,” IEEE Commun. Lett., vol. 7, no. 11, pp. 525–527, Nov. 2003.

替换策略-被动

- Neighbor Caching-LRU[2]：

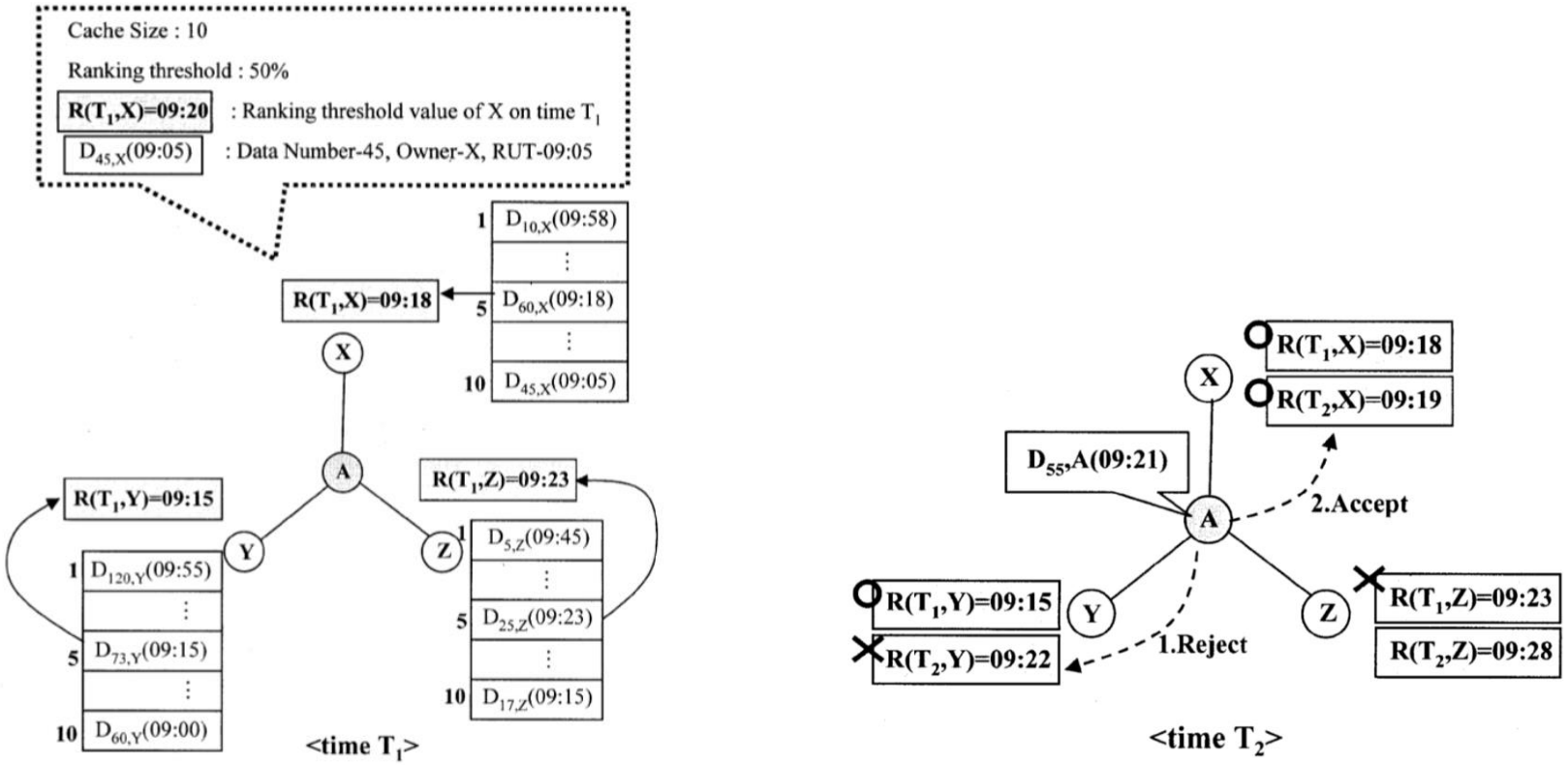


Fig. 1. Choosing the best neighbor according to the ranking based prediction.

[2] J. Cho, S. Oh, J. Kim, H. H. Lee, and J. Lee, “Neighbor caching in multi-hop wireless ad hoc networks,” IEEE Commun. Lett., vol. 7, no. 11, pp. 525–527, Nov. 2003.

替换策略-主动

- 主动边缘缓存策略主要关注四个指标，即流行度、上下文信息、社会因素和移动性信息。

- 基于流行度：如GreenDelivery[1]
- 基于上下文信息：通过观察连接到BS的用户的上下文信息来对内容进行主动缓存[2]。

[1] S. Zhou, J. Gong, Z. Zhou, W. Chen, and Z. Niu, “GreenDelivery: Proactive content caching and push with energy-harvesting-based small cells,” IEEE Commun. Mag., vol. 53, no. 4, pp. 142–149, Apr. 2015.
[2] S. Müller, O. Atan, M. V. D. Schaar, and A. Klein, “Context aware proactive content caching with service differentiation in wireless networks,” IEEE Trans. Wireless Commun., vol. 16, no. 2, pp. 1024–1036, Feb. 2017.

替换策略-主动

- 基于社会因素的[3]：将社会关系作为度量，在该模型中，定义了三个用户状态。
  - 易感状态(Susceptible State):用户没有收到任何关于内容的信息。
  - 感染状态(Informed State)，用户已经浏览了内容，并愿意通过在线社交网络与可能有共同兴趣的网友分享。
  - 不感染状态(Refractory State)，即用户已经查看了内容，但不会与其他人共享，也不会再次查看。

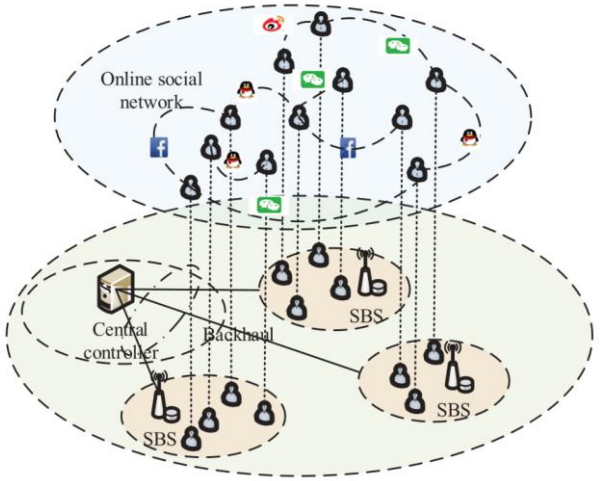


Fig. 1. The scenario of edge caching in social network.

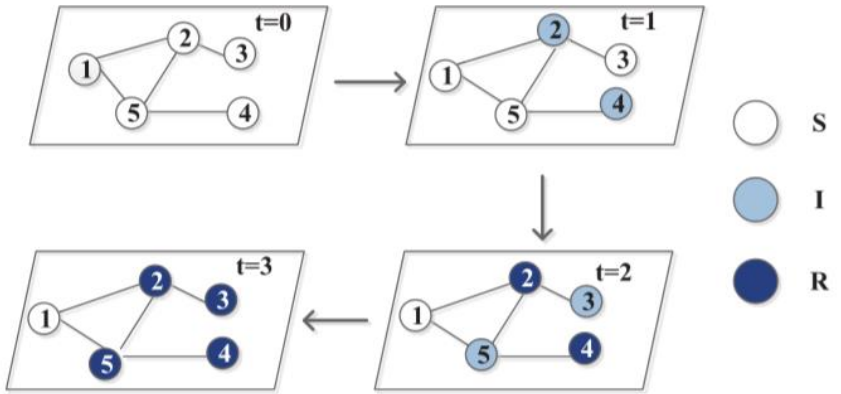


Fig. 2. The diagram of spreading process of a content.

[3] S. He, H. Tian, and X. Lyu, “Edge popularity prediction based on social-driven propagation dynamics,” IEEE Commun. Lett., vol. 21, no. 5, pp. 1027–1030, May 2017.

替换策略-主动

- 基于移动性信息：对于移动能力强、切换频繁的用户，在mmWare BSs中进行内容缓存[4]。

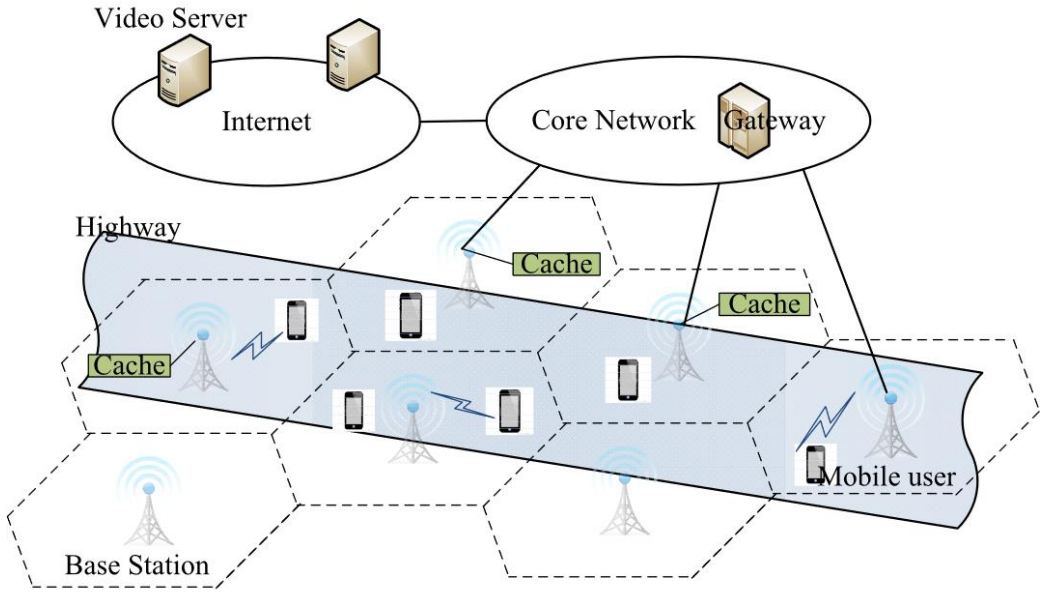


Fig. 1. Video Streaming Architecture with Cache Memory.

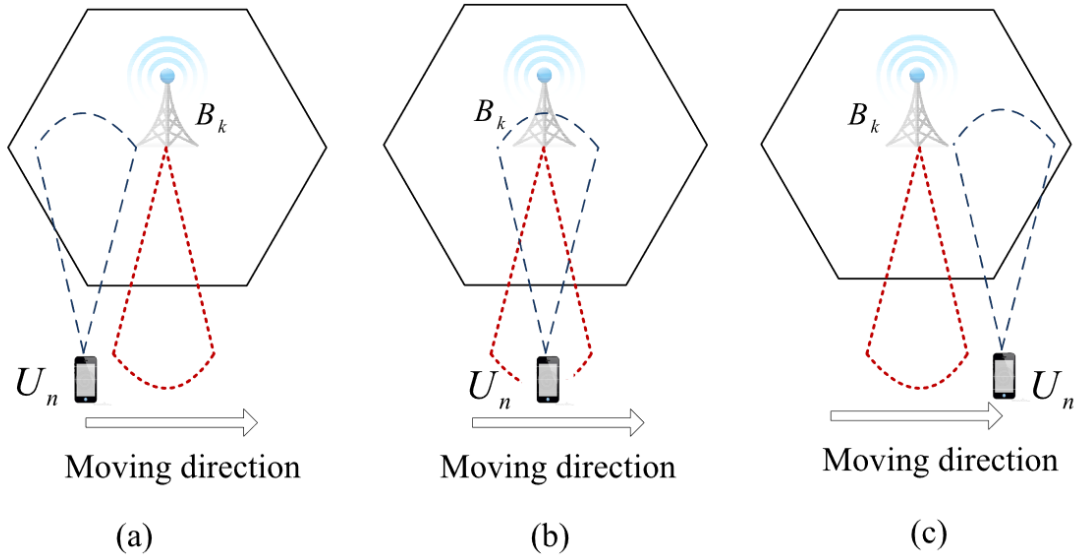


Fig. 2. Directional mmWave Connection with High Mobility.

[4] J. Qiao, Y. He, and X. S. Shen, “Proactive caching for mobile video streaming in millimeter wave 5G networks,” IEEE Trans. Wireless Commun., vol. 15, no. 10, pp. 7187–7198, Oct. 2016.

预取算法

- 以内容流行度为例
  - 静态分布：静态分布是指假设内容流行度已知或者服从简单特定的分布。

  如Zipf分布：

$$p_i = \frac{\frac{1}{f_i^{\gamma}}}{\sum_{i=1}^{F} \frac{1}{f_i^{\gamma}}}, \quad 1 \leqslant f_i \leqslant F$$

  另外，李敏提出 Zipf 模型并不适合描述视频请求概率[1]，并使用广延指数模型进行刻画，其概率密度函数为 ：

$$p(x)\,\mathrm{d}x = c\left(\frac{c^{c-1}}{x_0^{c}}\right)\exp\left\{-\left(\frac{x}{x_0}\right)^{c}\right\}\mathrm{d}x$$

[1] 朱敏，李俊． 视频点播中视频流行度的建模与分析［J］． 电子技术，2016(9) : 12． Zhu Ming，Li Jun． Modeling and analysis of video popularity in video-on-demand［J］． Electronic Technology， 2016(9) : 12．

预取算法

- 以内容流行度为例
  - 时变分布：其假设内容流行度随时间变化。

  ①相关性，例如在就餐过程中，先到客户的点餐会对后来的客户点餐产生影响.

  ②时变特性，例如流行的内容和新闻在刚出现时流行度低，接下来流行度急剧增加，然后流行度逐渐（ 或者急剧）降低.

  - 学习和预测流行度：采用机器学习或者矩阵分解的方法进行预测。

目录

非易失缓存管理

- 雾计算需要较大的主存容量来降低延迟，提高服务质量[1]

- DRAM
  - 缺点：高能耗，低密度；
  - 优点：高速的内存访问；
- PCM
  - 缺点：高延迟，有效的写次数；
  - 优点：高密度，良好的扩展性，低idle能耗，非易失；

TABLE I
COMPARISON OF THE DRAM AND THE PCM PERFORMANCES

Attributes	DRAM	PCM
Read latency	30–50 ns	50–100 ns
Write latency	30–100 ns	200–800 ns
Average energy	~0.1 nJ/b	~0.3 nJ/b
Idle power	~1.3 W/GB	~0.05 W/GB
Endurance	∞	10^8 for write

[1] Jia G, Han G, Wang H, et al. Cost aware cache replacement policy in shared last-level cache for hybrid memory based fog computing[J]. Enterprise Information Systems, 2018, 12(4): 435-451.

非易失缓存管理

- 在last-level cache处使用LRU策略提高hit ratio，能减少主存访问，提升混合内存的表现；
- 以往的工作也都是基于统一的内存(DRAM)平台进行的，他们的方法也都不适用于混合内存平台。
- 因为LLC miss ratio的减少不代表混合内存平台代价的减少(相比DRAM内存平台)。

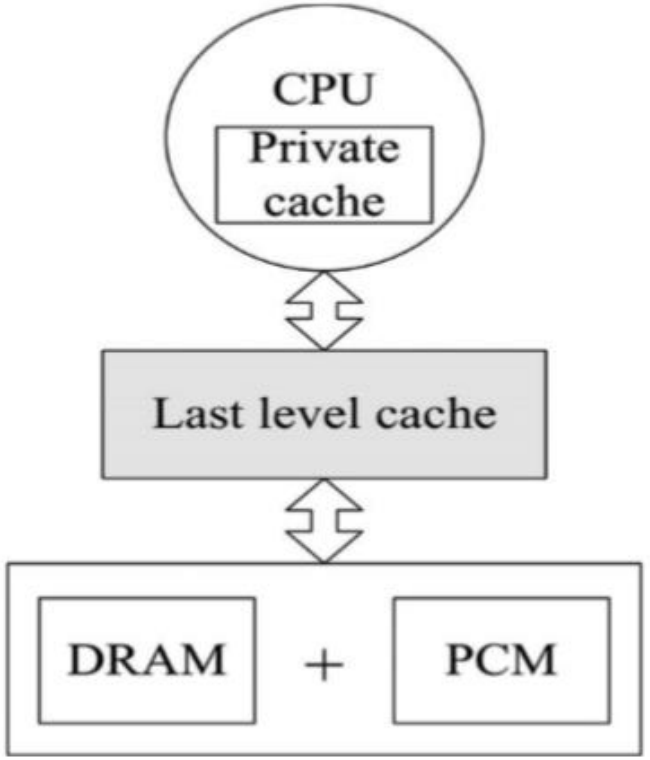


Figure 1. The hybrid memory architecture.

Jia G, Han G, Wang H, et al. Cost aware cache replacement policy in shared last-level cache for hybrid memory based fog computing[J]. Enterprise Information Systems, 2018, 12(4): 435-451.

非易失缓存管理

- LRU对于混合存储不再适用。

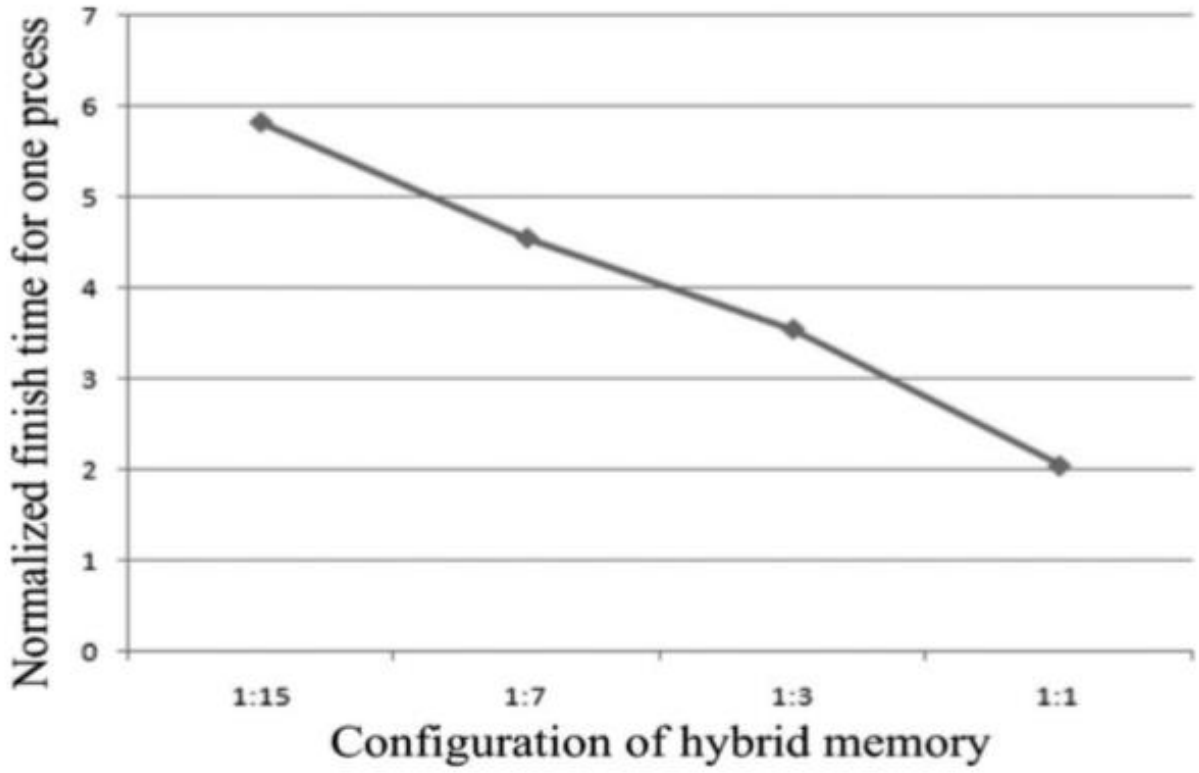


Figure 2. Normalized finish time under different configurations of hybrid memory using LRU.

非易失缓存管理

- 在DRAM和NVM的cache miss cost是不同的；
  - DRAM-only

$$AT_{memory} = H * T_{cache} + M * T_{memory}$$

  - Hybrid

$$AT_{hmemory} = H * T_{cache} + M_D * T_{memoryD} + M_N * T_{memoryN}$$

- 混合内存系统中的LLC替换策略影响cache性能。

非易失缓存管理

- 开销感知的缓存替换策略
- 三步走
  1. 将cache line区分为DRAM data和NVM data;
  2. 计算混合内存的miss cost;
     - miss cost受进程行为影响
  3. 确定data的插入位置。
     - NVM data插入在MRU位置;
     - DRAM data根据cost计算出插入位置。

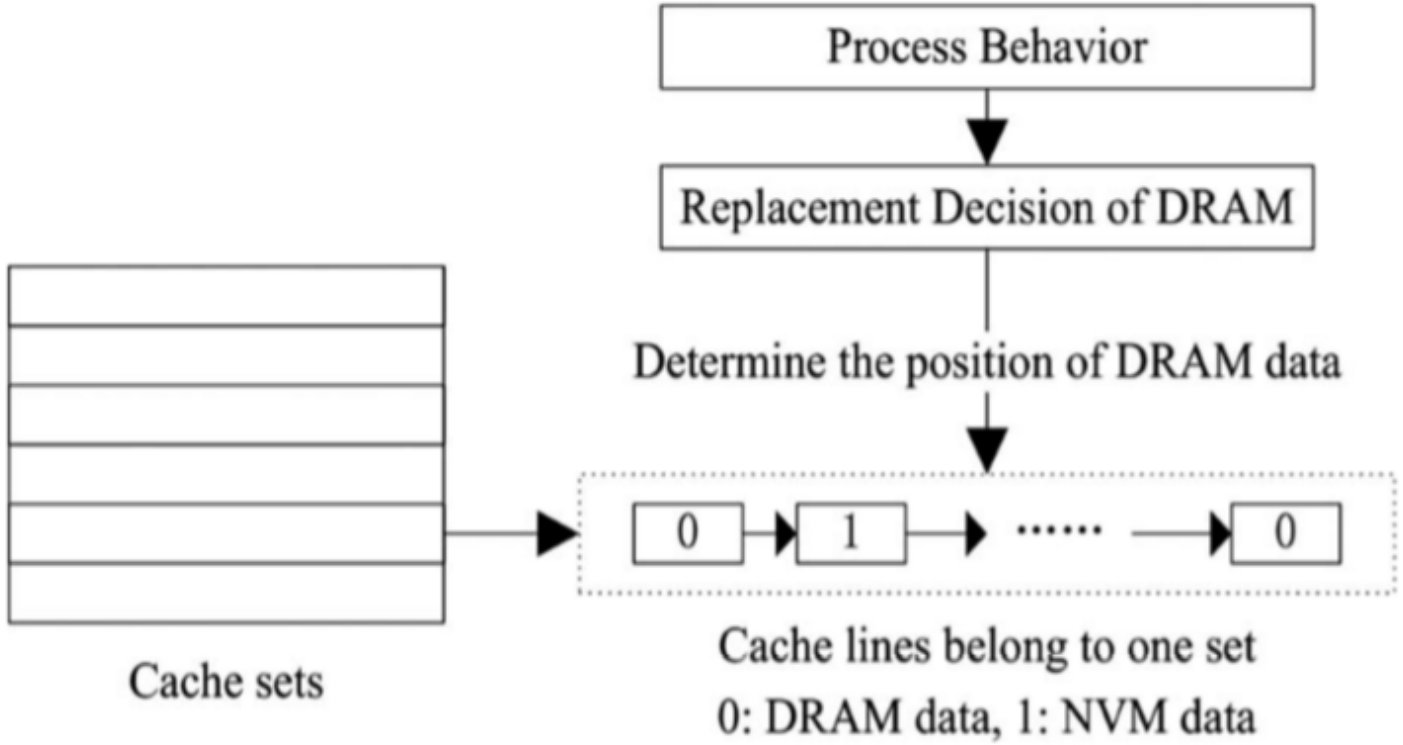


Figure 4. CACRP framework overview.

非易失缓存管理

S1.Cache line distinguishing mechanism

- 低地址归为DRAM data;
- 高地址归为PCM data;

Algorithm 1: Distinguishing algorithm

```
1: Input:
2: T: the tag of the candidate cache line
3: CL: the cache line size
4: S: the number of cache sets
5: CA = T * S * CL
6: Output:
7: 0 represents DRAM data
8: 1 represents PCM data
9: Begin
10: tem = CA%N;
11: if tem == 0
12:     return 0;
13. else
14:     return 1;
15: End
```

非易失缓存管理

S2.Replacement cost calculation

- DRAM Cost

$$C_{DRAM} = R_{rDRAM} * T_{rDRAM} + R_{wDRAM} * T_{wDRAM}$$

$$R_{rDRAM} + R_{wDRAM} = 1 \text{ and } T_{rDRAM} = T_{wDRAM}.$$

- PCM Cost

$$C_{NVM} = R_{rNVM} * T_{rNVM} + R_{wNVM} * T_{wNVM}$$

$$R_{rNVM} + R_{wNVM} = 1 \text{ and } T_{wNVM} > T_{rNVM}.$$

非易失缓存管理

S3.Adaptive position replacement policy

- Insertion Position Factor K

$$
\begin{aligned}
K &= C_{NVM}/C_{DRAM} \\
&= (R_{rNVM} * T_{rNVM} + R_{wNVM} * T_{wNVM})/(R_{rDRAM} * T_{rDRAM} + R_{wDRAM} * T_{wDRAM}) \\
&= (R_{rNVM} * T_{rNVM} + R_{wNVM} * T_{wNVM})/T_{rDRAM}
\end{aligned}
$$

- Insertion Position P

$$P = -2K/3 + 47/3$$

$$K > 1,\ 0 < P <\ = 15.$$

非易失缓存管理

S3.Adaptive position replacement policy

- 将cache line插入到LRU队列中；

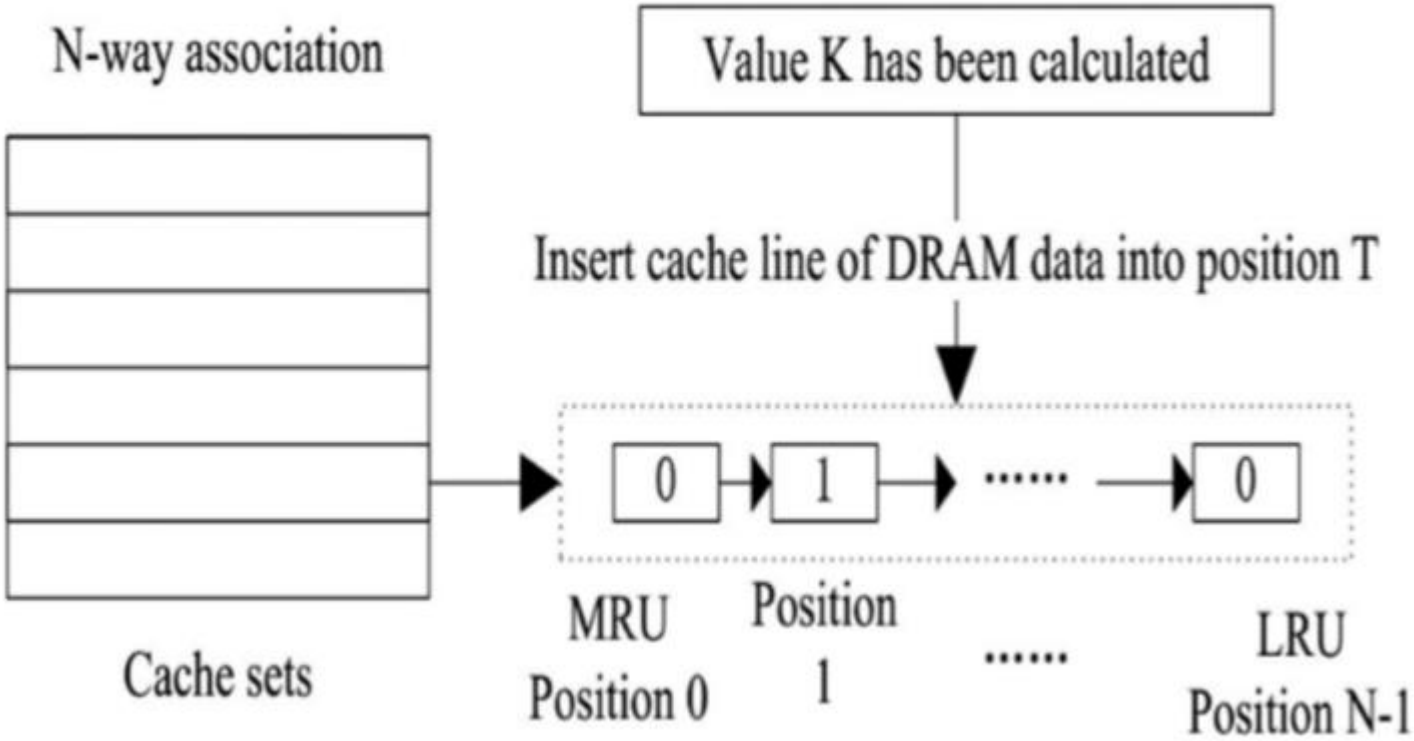


Figure 5. The adaptive position replacement policy.

目录

技术挑战和发展方向

- 隐私性和安全性
- 视频传输社交化
- 缓存技术商业化障碍
- 边缘缓存与非易失存储

谢　谢

敬请批评指正！