The cross-ISP traffic and performance tradeoff in VoD system

May 4, 2011

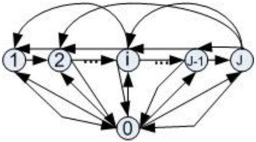
VoD System Model

- There are M ISPs: ISP 1, ISP 2, ISP 3,..., ISP M.
- The peer number and average peer upload bandwidth in different ISPs:
  - the peer number: We use the ON-OFF model. There are totally N_i peers (who have installed the software for VoD streaming) in ISP i. A part of N_i peers stays offline. The probability that a peer stays offline is π_0. $P(N_i^{off} = x) = C_{N_i}^{x} \pi_0^x (1-\pi_0)^{N_i - x}$.
  - The online peers are downloading chunks and uploading chunks in the system.
  - the average peer upload bandwidth in ISP i is U_i.

Chunk Demand in ISP i

- A total of J constant-length chunks to be shared: $C_1, C_2, ..., C_J$.
- At time slot T, peers download the chunks that will be played in time slot $T+1$.
- In ISP i, at time slot T, there are m_i^j peers downloading chunk j, and there are m_i^0 offline peers. $m_i^0 + m_i^1 + ... + m_i^J = N_i$. Say peers are in state $0, 1, 2, ..., J$ as peers are offline, downloading chunk 1, downloading chunk 2,...,downloading chunk J. The probability that peers are in some state is π_j for state j. $\sum_{j=0}^{J} \pi_j = 1$.
- The probability that there are $m_i^j (0 \le j \le J)$ peers in state j is $P(m_i^0, m_i^1, ..., m_i^J) = N_i! \frac{\pi_0^{m_i^0}}{m_i^0!} \cdots \frac{\pi_J^{m_i^J}}{m_i^J!}$.
- The probability that there are m_i^j peers downloading chunk $j (1 \le j \le J)$ is $P(m_i^j) = C_{N_i}^{m_i^j} \pi_j^{m_i^j} (1-\pi_j)^{N_i - m_i^j}$.

The chunk popularity in VoD system



Chain for user behavior.jpg

- User behaviors can be modeled by the state transition of peers. Based on the Markov Chain model for user behavior, we can calculate the equilibrium state distribution for peer state, $(\pi_0, \pi_1, ..., \pi_{J-1}, \pi_J)$. We can get the chunk popularity from this.
- User behaviors: Joining, Departures, Random seek.

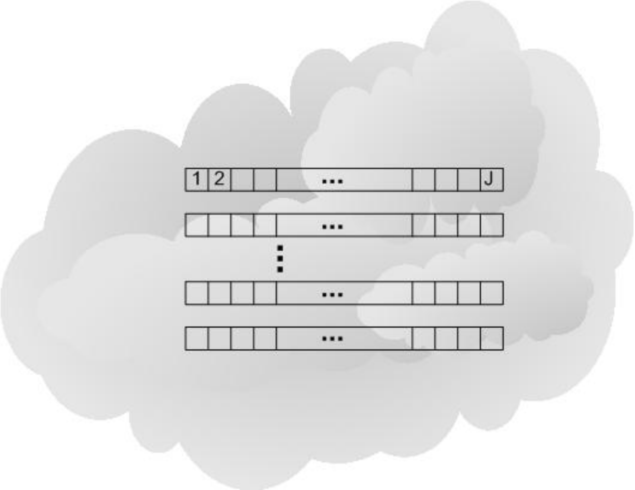
Define performance metrics

- The resource used to serve chunk j is from peers or from servers.
- At a time slot, m_i^j peers demand chunk j, we denote b_j as the number of peers that have cached the demanded chunk before (have played chunk j before), denote w_j as the copies of chunk j that online peers can upload.
- The needed server capacity to satisfy the demand for chunk j is $U_{sj} = max m_i^j - b_j - w_j, 0$. The total needed server capacity is $U_s = \sum_{j=1}^{J} U_{sj}$.
- If the server capacity is given as U_s, the probability of chunk missing for streaming is $P = \frac{\sum_{j=1}^{J} U_{sj} - U_s}{N_i - m_i^0}$.

chunk distribution in a peer' cache

- the chunk distribution in peers' cache relates to two aspects
  - a peer's sojourn time in VoD system
  - the chunk popularity
- How to calculate the chunk distribution
  - A video has J chunks, $C_1, C_2, \ldots, C_J$.
  - We assume that when a peer becomes offline from a video for sometime, the buffer that stores the chunks from the video is replaced by other videos' chunks.
  - For a peer, the probability for its ON duration of T time units is $P(T) = (1 - \pi_0)^T \cdot \pi_0$.
  - Assuming no prefetching, a peer stays online for time t, the chunk distribution in a peer online for t can be calculated based on the path of state transitions in time t. *e.g.* When a peer's state transits from state 3 to state 5 to state 6, it has chunk 3, 5, 6 in its buffer.
  - With the chunk popularity and the peers' sojourn time, we can calculate the chunk distribution in a peer's cache.

chunk distribution in a peer's cache



Chunk Supply

- Calculate the chunk supply from peers
  - relates to the chunk distribution in peers' cache and peer scheduling and chunk scheduling.

Next step work

- Based on the chunk distribution in peers' cache, and a specific peer selection and chunk scheduling, calculate the chunk supply w_j.
- Calculate the cross-ISP chunk demand and supply based on locality biased-overlay.