Proposition **DISCONT**
(Evaluation de systèmes de DIStribution de CONTenus)

Partenaires : **Institut Eurécom**, **Université de Montpellier II**, **Projet MISTRAL**

March 27, 2002

Objet : Cette proposition COLOR porte sur l'analyse et l'évaluation des performances de systèmes de distribution de contenus et, plus particulièrement, de serveurs vidéo et de systèmes de caches distribués.

Les organismes, équipes et personnes concernés par cette proposition COLOR sont les suivants :

- les départments Communications d'Entreprise et Multimedia de l'Institut Eurécom, avec comme partenaires les professeurs **Ernst Biersack** et **Keith W. Ross**.
- le LIRMM, UMR de l'Université de Montpellier II avec le CNRS, avec comme partenaire le professeur **Alain Jean-Marie**.
- le projet MISTRAL.

La durée de ce projet est de **1 an**.

Descriptif scientifique

Nous avons identifié deux thèmes de coopération, l'un porte sur l'analyse des performances d'un serveur vidéo, l'autre sur l'analyse d'un ensemble de caches distribués.

I – Performance evaluation of open-loop video distribution schemes for the case of bandwidth varying channels and user interactions

Content distribution and especially video distribution has been an important area of research for some time. One particular class of problem that has been studied is *open-loop* VOD distribution: Given that there are a large number of users who are interested in viewing the same video at approximately the same time, the requests of these clients are not satisfied one-by-one but in the following way:

- A video of length L is partitioned into N segments of length S_i, $i \in \{1, ..., N\}$ (L and S_i are expressed in units of time).
- Each segment $i, i \in \{1, ..., N\}$, is transmitted via broadcast at a constant rate d_i on a separate channel.
- A client who wants to play a video must receive the N segments by "listenig" to the different channels over which these segments are broadcast. The playout (consumption) of the video by a client can start as soon as the first segment has been entirely received. We call this instant T_p. To assure *continuity* in the playout (avoid starvation at the client), the rates d_i must be chosen sufficiently large to ensure that the client has entirely received a segment k no later than at time $T_p + \sum_{i=1}^{k-1} S_i$ (i.e. before the consumption of segment k starts)

There exist a large number of publications that investigate different ways of choosing the segment sizes and the rates of each segment [1]. However, there are two issues that have so far received no or very little attention:

- All the open-loop VOD distribution schemes assume that the channel for transmitting a segment i provides at all instances of time a bandwidth larger or equal to d_i. While this assumption is true for cable distribution networks, is does not hold for the transmission over the Internet, where the available bandwidth varies over time as a function of bottleneck bandwidth and the number of connections sharing that bottleneck bandwidth.

  We propose to model the Internet as a transmission channel where the available bandwidth can take a finite number of states $\{1, .., S\}$ (the distribution of the sejourn times is given, or at least its first moments) with values $b_j, j \in \{1, .., S\}$. We then want to determine for a segment i the minimal transmission rate $r_i, r_i > d_i$ such that segment i is received entirely before its playout must start. We plan to validate our model on the Internet.
- Traditional open-loop VOD distribution schemes do not support user interactions, i.e. a user can view a video only at its regular speed. Operations such as FF, Jump, Reverse, or Pause are not possible. Interactions where a user moves **forward**, such as FF or Jump, are particularly challenging to support since the consumption of all video data not yet visualized will occur earlier in time compared to the case where the client would view the video at its normal speed. Given that we have a characterization (distribution and its moments) of the duration of FF and Jump, we would like to compute for each a segment i the minimal transmission rate $r_i, r_i > d_i$ such that segment i is received entirely before its playout must start. The validity of the analytical results will be investigated through experiments to be conducted at Eurecom.

In terms of modeling, we will be using tools from regenerative or semi-regenerative process theory that seem to fit well with the problem description and, in particular, with the changing state of the communication channel.

II – Fluid models for shared Web caches

With a Web cache (also known as a proxy), a user first sees if a Web object is in a nearby cache; if the object is not in the cache, the cache retrieves the object from the Internet, stores a copy, and passes a copy to the requesting user. Web caches are used extensively by institutions (companies and universities) as well as ISPs.

Due to limited CPU, storage and bandwidth, a cache running on a single machine is often insufficient to meet the demand of a large set of users. One natural solution to this problem is for the provider (institution or ISP) to install many caches, each serving a different set of Web objects. The CARP protocol, now implemented in Microsoft and Netscape products, provides an elegant and decentralized means for coordinating the group of caches. The CARP protocol specifically addresses cache failures and growth in the set of caches [2].

In this research we are developing stochastic models for cooperative groups of caches. There are two random phenomena in these models: (1) object randomness, that is, times at which objects are requested and the size of objects; (2) cache up/down randomness, that is times at which caches go down, come up, or get added to the set.

Stochastic models that account for the microscopic behavior of both of these two types of randomness tend to be intractable by analytical means, and difficult for simulation analysis. However, by making the observation that the time scale of object requests and arrivals is orders of magnitude faster than the time scale of up/down dynamics, we may safely replace the microscopic randomness of requests/arrivals by a fluid model.

Specifically, we consider a model with N caches. At any instant of time let $N(t)$ be the number of caches that are up, and let $X(t)$ be the number of cached bytes in the aggregate set of caches. Each down cache comes up at rate lambda, and each up cache goes down at rate mu. The process $N(t)$ is the so-called *Engset process*, which is popular in teletraffic theory. Let (n, x) represent the current state of the system, where n is the number of up caches and x is the total fluid (number of document bytes) in the aggregate group of caches. When in state (n, x), we model the rate at which new fluid (new documents) enter the system by a deterministic function $f(n, x)$. Whenever there is a transition in $N(t)$, the fluid x is permitted to jump down. Let $g_u(n, x)$ be how much fluid leaves the system when there is an cache-up event; let $g_d(n, x)$ denote how much leaves the system when there is a cache-down event. As example, a good model for CARP has $f(n, x) = \sigma(1 - Z/X(t))$, $g_d(n, x) = x(n-1)/n$ and $g_u(n, x) = xn/(n+1)$.

In this research we will

- Model a variety of cooperative caching systems and determine appropriate f, g_u and g_d functions.
- Develop efficient simulation techniques for simulating these stochastic fluid models.
- Develop an analytical theory for tractable choices of f, g_u and g_d.

References

[1] A. Hu, "Video-on-Demand broadcasting protocols: A comprehensive study," *Proc. IEEE Incocom'2001 Conf.*, Vol. 1, pp. 508-517, Anchorage, Alaska, USA, April 2001.

[2] K. W. Ross, "Hash Routing for Collections of Shared Web Caches," *IEEE Network*, pp. 42-51, November/December 1997.