PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO

Rafael Silva Pereira

A Split&Merge Architecture for Distributed Video Processing in The Cloud

DISSERTAÇÃO DE MESTRADO

Dissertation presented to the Postgraduate Program in Informatics of the Departamento de Informática, PUC-Rio as partial fulfillment of the requirements for the degree of Mestre em Informática

Advisor: Prof. Karin Breitman

Rio de Janeiro, April 2011

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO

Rafael Silva Pereira

A Split&Merge Architecture for Distributed Video Processing in The Cloud

Thesis presented to the Postgraduate Program in Informatics, of the Departamento de Informática do Centro Técnico Científico da PUC-Rio, as partial fulfillment of the requirements for the degree of Mestre.

Prof. Karin Breitman
Advisor
Departamento de Informática – PUC-Rio

Prof. Fabio Kon
Departamento de Ciência da Computação – USP

Prof. Markus Endler
Departamento de Informática – PUC-Rio

Prof. José Engenio Leal
Coordenador Setorial do Centro
Técnico Científico - PUC-Rio

Rio de Janeiro, 18 de Abril de 2011

Rafael Silva Pereira

Graduated in Electronics and Computer Engineering at Universidade Federal do Rio de Janeiro – UFRJ in 2006. Media Engineering Specialist from Globo.com since 2006.

Resumo

Pereira, Silva Rafael; Breitman, Koogan Karin. **Uma Arquitetura de Split&Merge para Processamento Distribuido de Vídeo Baseado em Cloud.** Rio de Janeiro, 2010. 76p. Dissertação de Mestrado – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

O volume de dados existentes aumenta a cada dia, sendo que, armazenar, processar e transmistir esta informação se torna um grande desafio. O paradigma de Map Reduce, proposto por Dean e Ghemawat [10], é uma forma eficiente para o processamento de grandes volumes de dados utilizando um cluster de computadores e, mais recentemente, infraestruturas no Cloud. Entretanto, implementações tradicionais de Map Reduce não apresentam nem a flexibilidade (para escolher entre diferentes técnicas de codificação na etapa de map), nem o controle (capaz de especificar como organizar os resultados na etapa de reduce), necessários para o processamento de vídeos. Porém, com a proliferação de dispositivos capazes de reproduzir conteúdo em multimídia, e com o aumento da disponibilidade de banda, o consumo deste tipo de conteúdo é cada vez maior, o que mostra a necessidade de termos arquitetura eficientes para lidar com grandes volumes de dados, especificamente vídeos. A arquitetura de Split&Merge, proposta nesta dissertação, generaliza o paradigma de Map Reduce, fornecendo uma solução eficiente que contempla aspectos relevantes às aplicações de processamento intensivo de vídeo. Para validar a arquitetura proposta, são apresentados dois casos de uso onde a mesma foi implementada utilizando uma plataforma de Cloud.

Palavras-chave

Computação na Nuvem; Compressão de Vídeo; Sistemas Distribuídos; Arquiteturas Orientadas à Serviço.

Abstract

Pereira, Silva Rafael; Breitman, Koogan Karin (Advisor). **A Split&Merge Architecture for Distributed Video Processing in The Cloud.** Rio de Janeiro, 2010. 76p. MSc. Dissertation – Departamento de Informática, Pontifícia Universidade Católica do Rio de Janeiro.

The Map Reduce approach, proposed by Dean & Ghemawat [10], is an efficient way for processing very large datasets using a computer cluster and, more recently, cloud infrastructures. Traditional Map Reduce implementations, however, provide neither the necessary flexibility (to choose among different encoding techniques in the mapping stage) nor control (to specify how to organize results in the reducing stage), required to process video files. The Split & Merge tool, proposed in this thesis, generalizes the Map Reduce paradigm, and provides an efficient solution that contemplates relevant aspects of intense processing video applications.

Keywords
Cloud Computing; Video Compression; Distributed Systems; Service Oriented Architectures.

Table of Contents

List of Figures

List of Tables

Glossary

API - An application programming interface (API) is an interface implemented by a software program to enable interaction with other software, similar to the way a user interface facilitates interaction between humans and computers.

Codec - A codec is a device or computer program capable of encoding and/or decoding a digital data stream or signal. The word codec is a portmanteau (a blending of two or more words) of 'compressor-decompressor' or, more commonly, 'coder-decoder'.

DV - DV is a format for recording and playing back digital video. It was launched in 1995 with joint efforts of leading producers of video camera recorders.

DCT - "Discrete Cosine Transform", a mathematical transform that can provide aliasing cancellation and good frequency resolution, used in some codecs to convert the audio or video signal from the time domain to the frequency domain.

GOP - Group of Pictures, it starts with one I-frame and ends with the next I-frame, excluding the next I-frame.

HD - High definition. Usually used to describe any device capable of generating or displaying a signal with a resolution of at least 720 vertical lines (i.e. 720p).

HTML - HTML, which stands for HyperText Markup Language, is the predominant markup language for web pages.

HTML 5 - HTML5 is being developed as the next major revision of HTML. In particular, HTML5 adds many new syntactical features. These include the *<video>*, *<audio>*, and *<canvas>* elements, as well as the integration of SVG content. These features are designed to make it easy to include and handle multimedia and graphical content on the web without having to resort to proprietary plugins and APIs.

HTTP/HTTPS - HTTP is an application protocol, which defines how files on the World Wide Web are transferred. HTTPS (HTTP over SSL or HTTP Secure) is the use of Secure Socket Layer (SSL) or Transport Layer Security (TLS) as a sublayer under regular HTTP.

NFS - Network File System. Standard for accessing files on a remote computer appearing as a local volume.

OCR - Optical Character Recognition. The technology that allows computers to 'read' text from files, such as images.

RGB - An additive color model based on red (R), green (G), and blue (B) light. RGB is used by computers, televisions, and film recorders to display colors.

SD - Standard definition. Usually defined as a 480i signal (480 interlaced scan lines) presented 30 times per second.

SLA - A service level agreement (frequently abbreviated as SLA) is a part of a service contract where the level of service is formally defined. In practice, the term SLA is sometimes used to refer to the contracted delivery time (of the service) or performance.

UGC - User-generated content (UGC), also known as consumer-generated media (CGM) or user-created content (UCC), refers to various kinds of media content, publicly available, that are produced by end-users.