

# Caracterização da Transmissão de um Grande Evento Esportivo

Wagner Almeida<sup>1,3</sup>   Breno Santos<sup>1</sup>   Alex B. Vieira<sup>1</sup>  
Ítalo Cunha<sup>2</sup>   Jussara Almeida<sup>2</sup>

<sup>1</sup>Departamento de Ciência da Computação, Universidade Federal de Juiz de Fora

<sup>2</sup>Departamento de Ciência da Computação, Universidade Federal de Minas Gerais

<sup>3</sup>Instituto Federal do Sudeste de Minas Gerais

{wagner.almeida, breno.santos}@ice.ufjf.edu.br   {alex.borges}@ufjf.edu.br  
{cunha, jussara}@dcc.ufmg.br

**Abstract.** *The unavailability of multicast in the Internet limits the scalability of streaming popular content like sports events. To alleviate this problem, content providers use alternative approaches like peer-to-peer distribution and content distribution networks. Capacity planning is imperative in both cases and, if done incorrectly, may lead to high costs, performance problems, and frustrated users. In this paper we characterize traffic logs collected from video servers at one of the largest online content providers in South America that streamed FIFA's 2013 Confederations Cup live. We quantify the workload on the streaming servers as well as the quality of service experienced by end users. We show that streaming servers receive substantial load and that load can vary significantly between games, making capacity planning challenging.*

**Resumo.** *A falta de implantação de multicast na Internet gera problemas de escalabilidade para distribuição de conteúdo popular como transmissão de eventos esportivos. Isto leva à utilização de soluções alternativas como distribuição em redes par-a-par ou por redes de distribuição de conteúdo (CDNs). Em ambos os casos é necessário planejamento de capacidade, que se não for feito de maneira correta pode gerar erros de provisionamento que resultam em problemas de desempenho e usuários frustrados. Neste artigo apresentamos uma caracterização do tráfego da transmissão ao vivo da Copa das Confederações da FIFA de 2013 realizada por um dos maiores provedores de conteúdo multimídia pela Internet da América Latina. Nós quantificamos o impacto da transmissão de um evento de grande escala no provedor de conteúdo e a qualidade de serviço recebida pelos usuários finais. Mostramos também que os servidores de vídeo recebem carga substancial e que a carga varia significativamente entre jogos, o que dificulta o provisionamento de capacidade.*

## 1. Introdução

Muitos espectadores estão deixando sistemas tradicionais de televisão, como cabo e satélite, e optando por assistir seus programas favoritos sob demanda através de serviços de *streaming* como Netflix [Mahanti 2014]. Neste cenário, sistemas de *streaming* ao vivo vêm recebendo atenção substancial, com um número expressivo de aplicações comerciais emergindo. Entretanto, devido à falta de implantação de *multicast* na Internet, a

distribuição de conteúdo multimídia em larga escala é um desafio significativo. Diversos mecanismos como *streaming* adaptável, uso de redes par-a-par e redes de distribuição de conteúdo (CDNs) vêm sendo utilizados para tornar possível a distribuição de mídia ao vivo em larga escala.

No entanto, ainda são necessários grandes investimentos em infraestrutura—como planejamento de capacidade e provisionamento de recursos—para atender à demanda crescente por esse tipo de conteúdo. Alguns eventos atraem um número muito grande de espectadores, impondo aos sistemas dos provedores de conteúdo e à infraestrutura de rede uma carga de trabalho difícil de prever e que pode exigir mais recursos do que aqueles já provisionados.

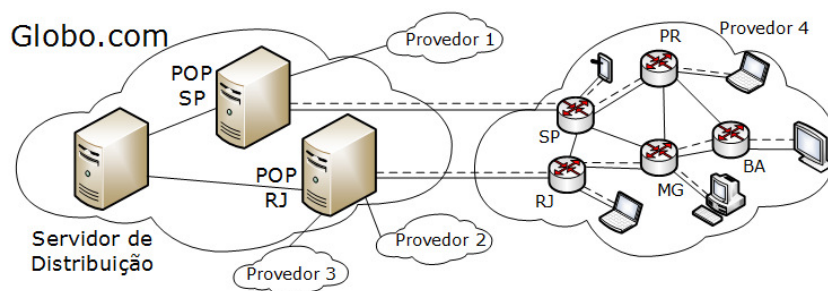
Entre as transmissões ao vivo com potencial para atrair grande interesse estão os eventos esportivos. Em particular, os jogos da Copa do Mundo (e seu torneio preparativo, a Copa das Confederações) atraíram a atenção do mundo todo e, devido à paixão local pelo esporte, em especial do Brasil. Espera-se que a transmissão dos jogos olímpicos de 2016 tenha popularidade similar.

Uma tarefa desafiadora no planejamento de capacidade de provedores de conteúdo é prever o volume de tráfego e provisionar banda de rede suficiente para eventos futuros. O volume de tráfego e provisionamento são diferentes para cada provedor, dependendo de questões como localização geográfica, parcerias de rede e número de clientes. As cargas também são diferentes para propriedades específicas do evento como horário e popularidade dos times. O provisionamento de recursos deve ainda considerar requisitos de aplicação, como baixa latência de distribuição.

Neste artigo, apresentamos uma caracterização da transmissão ao vivo de todos os jogos da Copa das Confederações, que é um grande evento preparatório para a Copa do Mundo da FIFA. Nosso trabalho é baseado em *logs* coletados nos servidores de vídeo de um dos maiores e mais conhecidos provedores de conteúdo da América Latina, o Globo.com. Nosso objetivo é entender qual é o impacto de um evento dessa magnitude na carga imposta aos servidores do provedor de conteúdo e na infra-estrutura de rede, bem como avaliar a qualidade de serviço entregue aos usuários.

Nossos resultados mostram que um evento como a Copa das Confederações tem potencial de atrair centenas de milhares de espectadores simultâneos. O tráfego gerado por transmissões de eventos ao vivo impõe desafios de infraestrutura, planejamento e previsão de carga até mesmo para um grande provedor; por exemplo, observamos que o provedor reajustou a qualidade da mídia disponível durante a transmissão de um jogo devido à alta carga. Além disso, pudemos observar que as regiões mais desenvolvidas do Brasil obtiveram maior qualidade na transmissão dos jogos. Isso pode ser relacionado à qualidade no acesso à Internet como um todo nessas regiões.

Nossa primeira contribuição é apresentar uma caracterização profunda do *streaming* e da carga de trabalho à qual o provedor de conteúdo é submetido durante um grande evento. Nessa linha, caracterizamos o número de usuários de cada jogo, a taxa de mídia distribuída e a qualidade da mídia. Nossa segunda contribuição é caracterizar a qualidade de serviço recebida pelos clientes e relacioná-la a dados geográficos e sócio-indicadores. Acreditamos que nossos resultados podem ser úteis para provedores de conteúdo ajudando a direcionar decisões de provisionamento de infra-estrutura e planejamento da capacidade



**Figura 1. Exemplo de *anycast***

das redes dos provedores de conteúdo visando o *streaming* de grandes eventos.

## 2. Descrição do Sistema e Conjunto de Dados

Nesta seção analisamos a arquitetura e principais componentes do sistema de transmissão de vídeo ao vivo do Globo.com, bem como o conjunto de dados. A maioria das informações dos *logs* aqui descritas foram fornecidas pelo próprio provedor de conteúdo.

### 2.1. Infra-estrutura de Distribuição de Conteúdo

O sistema de transmissão de vídeo ao vivo do Globo.com usa *anycast* [Cesario 2012], uma técnica de engenharia de tráfego onde um prefixo IP é anunciado a partir de múltiplos locais. A rede então propaga os anúncios e decide livremente para qual localização encaminhar os dados, considerando a configuração dos protocolos de roteamento e possivelmente a disponibilidade do servidor anunciante [Katabi and Wroclawski 2000].

Como mostramos na figura 1, a infra-estrutura de *streaming* ao vivo possui dois pontos de distribuição em duas cidades pólo do Brasil, Rio de Janeiro e São Paulo. Em cada um desses pontos de distribuição, o provedor de conteúdo é conectado a um ponto de troca de tráfego (PTT) local e a várias redes comerciais. As requisições recebidas em um ponto de distribuição são balanceadas entre os diversos servidores disponíveis em cada um dos pontos de distribuição.

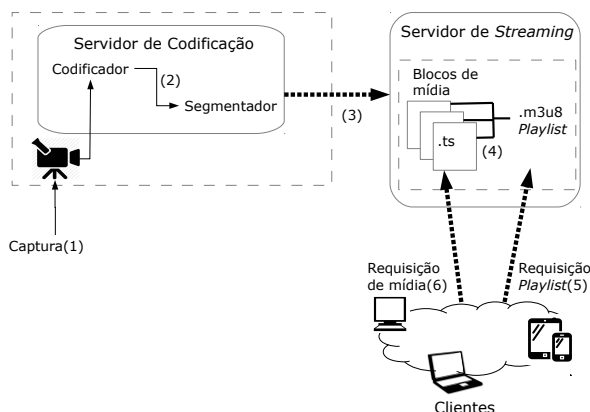
O sistema de transmissão de vídeo do Globo.com já foi utilizado para transmissão de vários eventos de grande escala. Os servidores do Globo.com executam o nginx <sup>1</sup>, um servidor Web de alto desempenho.

### 2.2. Interação Cliente-Servidor

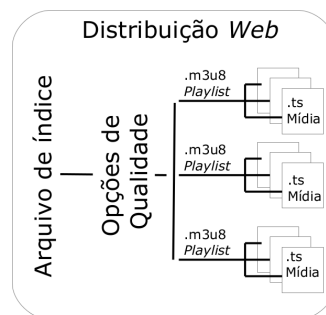
Atualmente, a maioria dos servidores de *streaming* de vídeo, como os servidores do Globo.com, YouTube e Netflix, entregam conteúdo usando HTTP. O uso de HTTP traz consigo a flexibilidade inerente desse ambiente e supera limitações de sistemas de *streaming* em P2P. Por exemplo, transmissão por HTTP requer somente um *browser* padrão para visualizar conteúdos de vídeo, sem necessidade de *software* adicional. Além disso, não é necessário manter portas adicionais TCP ou UDP abertas em *firewalls* e NATs.

Tais sistemas podem disponibilizar múltiplas taxas de codificação com diferentes níveis de qualidade usando *Dynamic Adaptive Streaming over HTTP* (DASH [Stockhammer 2011]). O DASH funciona dividindo o conteúdo de mídia em sequências de segmen-

<sup>1</sup><http://nginx.org/>



**Figura 2. Codificador e segmentador de vídeo**



**Figura 3. Múltiplas qualidades**

tos com taxas de codificação diferentes, de forma que cada cliente pode ajustar continuamente suas requisições de acordo com a estimativa local de disponibilidade de banda.

A transmissão de um evento ao vivo pela Internet não ocorre estritamente em tempo real. É possível observar na figura 2 que é necessário um pré-processamento envolvendo a codificação dos dados de som e imagem. A partir dessa codificação, o DASH divide o arquivo de mídia em múltiplos segmentos de diversas qualidades e gera arquivos de índice em formato *.m3u8*, também chamados de listas de reprodução (figura 3).

Para receber o conteúdo, clientes primeiro solicitam a lista de reprodução referente a ele e então solicitam segmentos da mídia. Cada vez que um segmento é recebido, o cliente pode acessar o servidor e fazer uma nova requisição.

### 2.3. Conjunto de Dados

Nosso conjunto de dados é composto de arquivos de *log* dos servidores nginx do Globo.com coletados durante transmissão ao vivo de jogos da Copa das Confederações 2013.

A Copa das Confederações é um torneio de futebol preparatório para a Copa do Mundo e foi realizado no Brasil entre 15 e 30 de Junho de 2013. Oito seleções estiveram presentes no torneio. A competição foi dividida em duas fases. Na primeira fase as oito equipes foram divididas em dois grupos. Na segunda as duas equipes de cada grupo com melhor desempenho se classificaram para uma disputa de semi-final seguida pela disputa de terceiro lugar e pela final.

O evento foi amplamente transmitido por redes de televisão aberta, por emissoras de rádio e pela Internet. Durante o torneio aconteceram 16 partidas. Foram coletados dados durante todos os dias em que ocorreram jogos desse torneio. Entretanto, focamos nossas análises no período entre uma hora antes e uma hora depois da ocorrência de cada jogo. Embora haja coleta de dados fora desse período, o volume de tráfego e o número de clientes é insignificante.

A tabela 1 resume nosso conjunto de dados. Analisamos todas as partidas e verificamos que elas receberam entre aproximadamente 27.000 e 465.000 IPs únicos. A

equipe técnica do Globo.com nos informou que o identificador de sessão foi desabilitado nos servidores nginx devido a problemas de desempenho na maioria dos jogos. Como não é possível identificar sessões de usuários através dos *logs* em todos os jogos (i.e., identificar quais requisições correspondem à exibição de um vídeo para um usuário), consideramos que cada IP seja referente a um cliente. Notamos que esta abordagem pode subestimar a quantidade de usuários assistindo às transmissões em redes com NAT.

A tabela 1 mostra resultados agregados para todo o período monitorado de cada jogo. Além disso, durante a transmissão de um único jogo, observamos até 180.000 espectadores simultâneos. Observamos também mais de 115 Terabytes transferidos durante uma partida e picos de até 350 GB/s. Até onde sabemos, esta quantidade de clientes e volume de tráfego são mais expressivos que os dados analisados anteriormente em outros trabalhos relacionados [Marfia et al. 2007, Jiangchuan Liu and Zhang 2008, Hei et al. 2007, Erman and Ramakrishnan 2013, Shafiq et al. 2013].

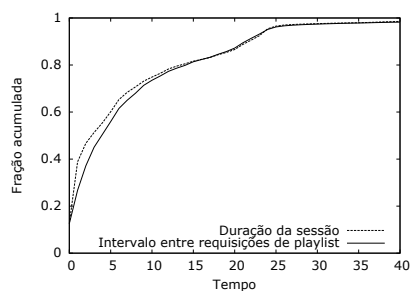
**Tabela 1. Resumo dos *logs* da Transmissão da Copa das Confederações**

Data	Hora	Jogo	IPs Únicos (Milhares)		
			RJ	SP	Total
15/06 - Sáb	16h	Brasil x Japão	93	149	242
16/06 - Dom	16h	México x Itália	38	84	122
16/06 - Dom	19h	Espanha x Uruguai	33	82	115
17/06 - Seg	16h	Taiti x Nigéria	72	114	187
19/06 - Qua	16h	Brasil x México	157	309	465
19/06 - Qua	19h	Itália x Japão	70	112	181
20/06 - Qui	16h	Espanha x Taiti	94	173	267
20/06 - Qui	19h	Nigéria x Uruguai	37	68	105
22/06 - Sáb	16h	Japão x México	10	17	27
22/06 - Sáb	16h	Brasil x Itália	61	110	172
23/06 - Dom	16h	Nigéria x Espanha	33	55	88
23/06 - Dom	16h	Uruguai x Taiti	32	55	87
26/06 - Qua	16h	Brasil x Uruguai	147	293	440
27/06 - Qui	16h	Espanha x Itália	156	230	387
30/06 - Dom	13h	Uruguai x Itália	72	119	190
30/06 - Dom	19h	Brasil x Espanha	126	214	340

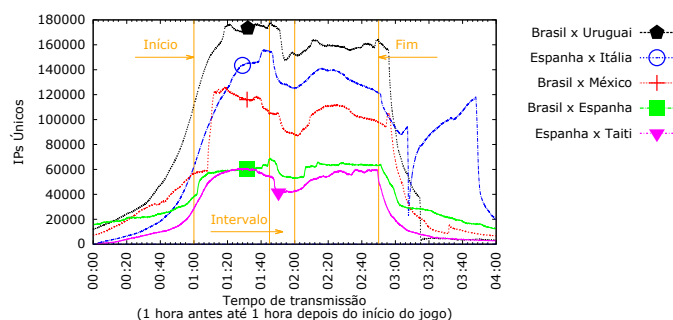
### 3. Caracterização da Carga nos Servidores

Nesta seção, apresentamos os resultados da caracterização da carga de trabalho imposta aos servidores durante as transmissões ao vivo dos jogos da Copa das Confederações. Para analisar a carga nos servidores, caracterizamos a quantidade de conexões atendidas, a taxa de transmissão média e o volume total de dados trafegados pelos servidores.

Analisamos os arquivos de *logs* de todos os jogos. Apenas os quatro primeiros exibiam identificadores de sessão. A figura 4 compara a distribuição acumulada da duração de sessão em um jogo com identificador e a distribuição acumulada do tempo entre chegadas de requisições para listas de reprodução em um jogo sem identificador de sessão. Observamos que ambos os tempos são aproximadamente uniformemente distribuído e menores que 40 segundos em 99% das requisições. Como clientes fazem requisições de listas de reprodução na escala de segundos, esperamos que a carga no servidor e qualidade da mídia recebida pelo cliente não variem significativamente num intervalo de 40 segundos. Além disso, esperamos que cada cliente ativo troque informações com o servidor pelo menos uma vez num intervalo de 40 segundos, e.g., para requisitar a próxima lista de reprodução que contém os próximos segmentos de mídia da transmissão. Usamos essas observações para definir que sessões de usuários terminam após 40 segundos



**Figura 4. Distribuição do tempo entre requisições de listas de reprodução e duração da sessão**



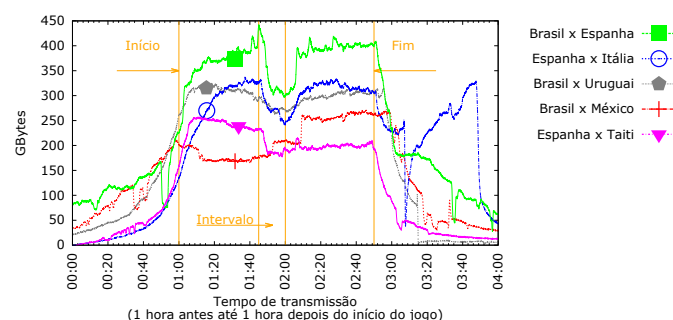
**Figura 5. Cinco partidas com maiores pico de acessos simultâneos**

sem requisição de listas de reprodução. Por essa razão nós calculamos métricas sobre o estado do sistema num instante de tempo agregando dados em intervalos de 40 segundos (*snapshots*).

Nós categorizamos os jogos de acordo com o pico de carga, calculado como o máximo de sessões de clientes simultâneas (num intervalo de 40 segundos) durante cada jogo. A figura 5 mostra os cinco jogos com os maiores picos de sessões simultâneas. Destacamos na imagem o início, o intervalo e o final das partidas (já considerando períodos de acréscimo típicos e excluindo partidas que tiveram disputa de pênaltis). Dessas cinco partidas, apenas Brasil e Espanha, que fizeram a final do torneio, se enfrentaram em um domingo; todos os outros jogos aconteceram em dias de semana entre 16:00 e 18:00, conforme mostrado na tabela 1.

É possível perceber ainda um maior interesse em jogos do Brasil, o anfitrião e equipe local, e em jogos da Espanha, a então primeira colocada no *ranking* FIFA. Brasil e Espanha tiveram média de 332.000 e 240.000 IPs únicos por partida, respectivamente. Além disso, todos os jogos entre os cinco com maiores números de espectadores envolveram pelo menos uma das duas equipes.

Uma tendência interessante a ser observada é a queda em número de usuários que ocorre entre o primeiro e o segundo tempo. Em todos os cinco jogos a média de espectadores cresce até os 20 primeiros minutos de partida e se mantém relativamente estável até o intervalo do jogo, aos 45 minutos. Da mesma forma, no segundo tempo, a média de usuários cresce nos primeiros 20 minutos e se mantém estável até o final. Por essa razão consideramos que os períodos entre 20 e 45 minutos de cada tempo da partida são os períodos de carga mais estável.



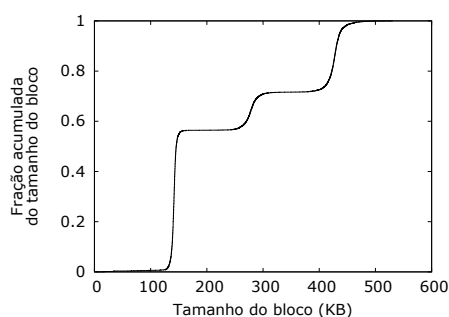
**Figura 6. Cinco partidas com maiores picos de volume de transmissão de dados**

Nos três jogos de maior carga, há uma queda de, em média, 9% entre o número de clientes conectados no período de pico do primeiro tempo e o número de clientes conectados no pico do segundo tempo. Esse foi o comportamento observado na maioria dos jogos. O jogo entre Brasil e Espanha apresentou um crescimento de aproximadamente 1,8% entre o primeiro e o segundo tempo. Observamos alguns jogos com esse comportamento, e podemos restringir a variação normal no número de clientes do primeiro para o segundo tempo nas partidas de média e alta carga dentro do intervalo de  $-10\%$  a  $+2\%$ . Vale a pena destacar ainda que o jogo entre Espanha e Itália foi o único entre os cinco de maior volume no qual aconteceram disputas de pênaltis após a duração regulamentar da partida. Por isso, também foi o único a apresentar uma grande quantidade de acessos na última hora do período analisado.

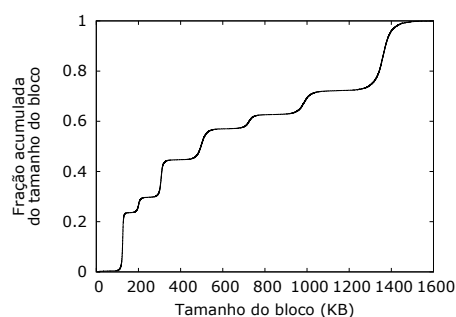
Analizamos também a carga através do volume de dados trafegados. A figura 6 mostra o volume de tráfego para os cinco jogos com maiores picos de volume de dados transferidos. Observamos que, durante o período de transmissão das partidas mostradas na figura 6, incluindo os intervalos, o tráfego foi bastante intenso. Em particular, a taxa de transferência se aproximou de 450 GB/s e nunca foi menor que 150 GB/s.

Note que a partida com maior volume de tráfego de dados (entre Brasil e Espanha) não foi a mesma partida com maior quantidade de clientes (entre Brasil e Uruguai). Isso aconteceu porque as opções disponíveis de qualidade do streaming de vídeo foram reduzidas. Com vídeos de menor qualidade, o provedor de conteúdo conseguiu atender à demanda não prevista de clientes. A partida entre Brasil e México foi a única que apresentou formato de curva diferente se compararmos as figuras 5 e 6. Apesar do número de clientes no segundo tempo de jogo da partida ter sido menor do que no primeiro, o volume de dados trafegados foi maior.

Estes resultados indicam que a taxa de transmissão varia muito entre jogos e durante um mesmo jogo. Para analisar a qualidade de vídeo e áudio fornecidos pelo Globo.com observamos o tamanho médio das requisições de segmentos de mídia de todas as partidas. Embora a distribuição geral, tenha sido omitida, nós analisamos e mostramos os resultados de alguns jogos específicos a seguir. No geral, observamos que há uma grande variação no tamanho médio de segmentos entre os jogos. Por exemplo, nos dois jogos com maior pico de acessos simultâneos, o tamanho dos segmentos é menor, quando comparado a outros jogos. Consequentemente, a qualidade da transmissão nesses dois eventos foi menor.



**Figura 7. Distribuição do tamanho dos segmentos de mídia do jogo entre Brasil e Uruguai**



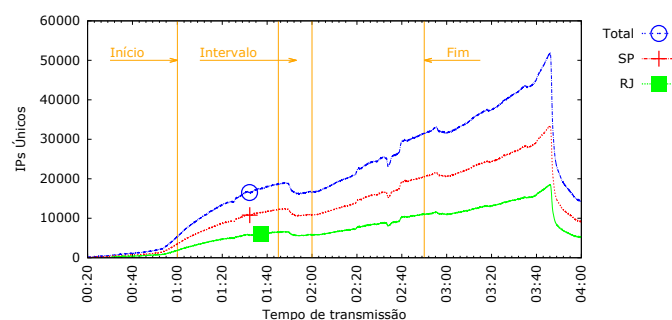
**Figura 8. Distribuição do tamanho dos segmentos de mídia do jogo entre Brasil e Espanha**

A figura 7 mostra a distribuição do tamanho dos segmentos de mídia transmitidos pelo servidor durante o jogo com maior número de espectadores (entre Brasil e Uruguai). Por essa figura, percebemos que há três platôs bem visíveis, provavelmente correspondendo a três taxas de codificação e níveis de qualidade distintos. Nesse dia, a transmissão com melhor resolução tinha blocos menores que 600 KB. Em contrapartida, no dia do jogo com maior volume de dados trafegados (entre Brasil e Espanha, figura 8) a qualidade do vídeo é maior, com segmentos de até 1600 KB. A figura 8 mostra que, além de maior taxa de codificação, há um número maior de taxas de codificação disponíveis. Devido a variações de desempenho entre diferentes *codecs* de vídeo, é impossível precisar exatamente qual é a qualidade da mídia que está sendo recebida pelos clientes com base somente no tamanho dos segmentos da mídia. Entretanto, como os tempos entre requisições de segmentos sofrem pouca variação (omitido), sabemos que a taxa de transmissão é proporcional ao tamanho dos segmentos.

O tamanho dos segmentos, por sua vez, é função da banda disponível no servidor e da banda disponível na rede do cliente. Esclarecendo, o servidor oferece várias opções de qualidade de mídia e tenta servir sempre a melhor possível. Se não há banda disponível, o servidor deve reduzir a qualidade da mídia e consequentemente o tamanho do bloco de vídeo. Por parte dos clientes, se há disponibilidade de banda na rede, a qualidade de mídia solicitada via DASH é maior, recebendo assim blocos de arquivos maiores.

Além das cinco partidas com maior carga, outra partida que nos chamou a atenção foi a realizada entre Uruguai e Itália, mostrada na figura 9. Essa partida, ao contrário das demais, apresentou crescimento da carga ao longo de toda a duração do jogo. Em uma análise mais profunda, dividimos o tráfego durante a transmissão desta partida recebido pelos dois pontos de distribuição do Globo.com. Nosso intuito foi de verificar se houve alguma anomalia em algum ponto de distribuição que pudesse impactar na análise como um todo. No entanto, o comportamento foi o mesmo nos dois pontos de distribuição de mídia ao vivo. Ressaltamos que essa partida foi a única realizada em um domingo às 13 horas. Portanto, conjecturamos que seu comportamento anormal se deve a essa diferença no horário ou dia de transmissão. Infelizmente, devido à curta duração e poucas partidas realizadas durante o campeonato, não há parâmetros de comparação que nos permitam confirmar essa especulação. Notamos que no jogo entre Uruguai e Itália houve disputa de pênaltis, por isso a carga no servidor continua aumentando após o término do jogo.





**Figura 9. Quantidade de usuários total e por localização do servidor da partida entre Uruguai e Itália**

Ainda na figura 9 mostramos que o número de clientes conectados aos servidores localizados no Rio de Janeiro é significativamente menor do que nos servidores localizados em São Paulo. Essa diferença aconteceu em todas as partidas, variando entre 50% e 60% de carga extra em São Paulo. É importante observar essa discrepância, uma vez que no planejamento atual do sistema, informado pelo Globo.com, a quantidade de banda disponível nos pontos de distribuição é praticamente a mesma.

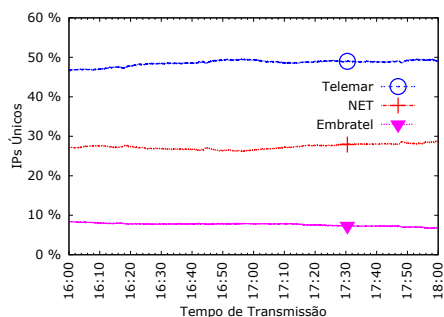
#### 4. Caracterização dos Clientes

Nesta seção caracterizamos a distribuição dos clientes entre os principais provedores de acesso à Internet no Brasil, correlacionando isto com a qualidade da mídia recebida, região geográfica e indicadores sócio-econômicos. Procuramos analisar o comportamento da rede e a distribuição dos usuários por provedores de acesso à Internet durante a transmissão das partidas. Para isso observamos os principais sistemas autônomos (ASes) correspondentes aos endereços IP dos espectadores. Para mapear endereços IP em sistemas autônomos usamos a base GeoIP da Maxmind <sup>2</sup>.

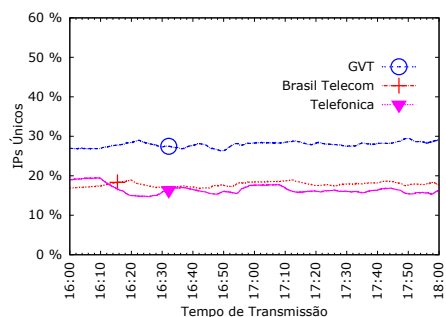
As figuras 10 e 11 mostram a quantidade de clientes nos três provedores com maior número de clientes assistindo a partida entre Brasil e Uruguai. Dividimos as figuras entre clientes conectados aos servidores do Rio de Janeiro e São Paulo, respectivamente. É possível observar que no Rio de Janeiro há um amplo domínio da Telemar, com cerca de 45% dos clientes. Em São Paulo a divisão é mais próxima, mas a Global Village Telecom (GVT) lidera com cerca de 27% dos clientes. Analisando os dados como um todo verificamos que, em São Paulo, as três empresas responsáveis pelos sistemas autônomos mostrados atendem a cerca de 70% de todo o tráfego da região. No Rio de Janeiro, esse percentual é ainda maior, chegando a 80%. Essa concentração de acesso reflete uma grande dependência de poucos provedores de acesso. Caso haja falha em algum dos seis ISPs citados neste trabalho, a provedora de conteúdo perderia uma quantidade relevante de seus clientes. Em um cenário ainda pior, falhas em mais de um desses provedores poderiam levar à indisponibilidade regional generalizada.

Com intuito de analisar a qualidade da mídia recebida pelos clientes dos principais provedores, analisamos a partida entre Brasil e Espanha, que foi a que apresentou maior volume de dados trafegados e uma variação maior de taxas de codificação da mídia (figura 8). Os resultados dessa análise podem ser vistos na figura 12. É possível perceber

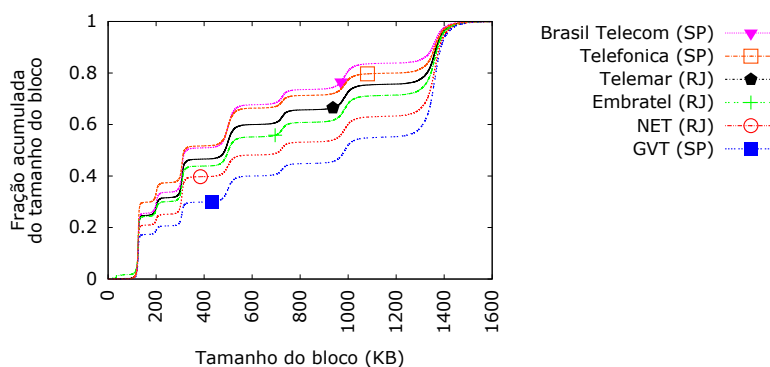
<sup>2</sup><http://www.maxmind.com/>



**Figura 10. Três ASs com maior número de usuários nos servidores do Rio de Janeiro**



**Figura 11. Três ASs com maior número de usuários nos servidores de São Paulo**



**Figura 12. Distribuição do tamanho dos segmentos de mídia por provedores no jogo Brasil e Espanha**

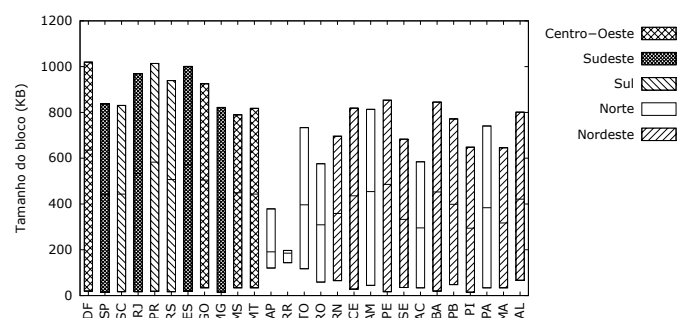
que GVT, NET e Embratel, em geral, recebem blocos maiores de segmentos de mídia se comparados a Telemar, Brasil Telecom e Telefônica. Isso parece indicar que os clientes dos primeiros provedores recebem um conteúdo de melhor qualidade do que os clientes dos outros provedores. Note que no Rio de Janeiro, embora a Telemar agrupe a maior quantidade de clientes, a qualidade recebida é inferior à dos clientes atendidos por Embratel e NET. Em São Paulo estão os provedores com maior e menor média de tamanho do bloco, respectivamente GVT, com média de aproximadamente 875 KB e Brasil Telecom, com média de aproximadamente 560 KB. Apesar de não termos informações para tirar conclusões sobre estas observações, elas motivam estudos sobre diversidade de rotas e provisionamento de capacidade nas redes dos provedores de acesso brasileiros.

#### 4.1. Caracterização Regional

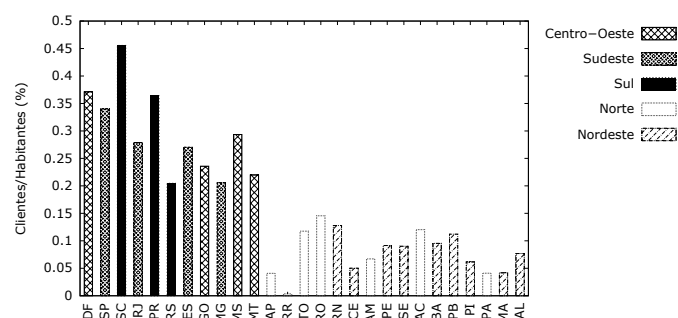
Encerramos nossa caracterização apresentando um mapa regional da qualidade da mídia distribuída para diferentes regiões do Brasil. Usamos as bases de dados GeoIP e GeoIP-City da Maxmind para identificar em quais regiões estavam localizados os IPs dos clientes do Globo.com. Também utilizamos dados de população e índice de desenvolvimento humano (IDH) disponibilizados pelo IBGE <sup>3</sup>.

A figura 13 mostra a mediana e os quartis do tamanho dos blocos de mídia recebidos por clientes agrupados em cada estado do Brasil. Ordenamos os estados no eixo  $x$

<sup>3</sup><http://www.ibge.gov.br/estadosat/>



**Figura 13. Tamanho do bloco de mídia por estado ordenado por IDH.**



**Figura 14. Relação Espectadores/habitantes por estado ordenado por idh**

pelo IDH em ordem decrescente. Observamos que à exceção do estado de Roraima, do qual a amostra de IPs foi muito pequena, todas as regiões apresentaram grande variação, mas nós observamos uma tendência geral de regiões com maior IDH terem maiores tamanhos de blocos. Isso é um indicador de que o acesso à Internet em banda larga é de melhor qualidade nessas regiões. A região Nordeste apresenta distribuição heterogênea, tendo estados com variados perfis de qualidade. Por fim, a região Norte é a que possui indicadores de qualidade de acesso em banda larga mais baixos.

Usamos o número de espectadores para tentar estimar a penetração da banda larga no Brasil. A figura 14 mostra uma relação percentual entre espectadores e habitantes de cada estado. Novamente ordenamos os estados no eixo  $x$  pelo IDH em ordem decrescente. Verificamos maior relação espectadores/habitantes nos estados das regiões Centro-Oeste, Sul e Sudeste bem como forte relação com o IDH. Essa relação também é observada na pesquisa nacional por amostra de domicílios (PNAD), realizada pelo IBGE em 2011<sup>4</sup> e pode indicar maior penetração de acesso à banda larga nessas regiões.

Aprofundamos a caracterização regional focando em cidades de três estados da região sudeste que apresentaram médias de tráfegos distintas entre suas capitais e algumas cidades pólo. A tabela 2 mostra os resultados dessa análise. Em geral, as cidades com mais habitantes receberam blocos maiores do que as cidades com menos. Observamos, nos estados do Rio de Janeiro e São Paulo, uma grande concentração de clientes nas capitais. Acreditamos que isso se deve à extensão das regiões metropolitanas dessas cidades. O estado de Minas Gerais apresentou distribuição geográfica dos clientes menos discrepante, com menor concentração na capital. Embora logicamente as cidades com

<sup>4</sup><http://www.ibge.gov.br/home/estatistica/populacao/acessoainternet2011/default.shtm>

**Tabela 2. Análise do tráfego por cidade de MG, RJ e SP**

Estado	Cidade	População	IPs únicos	IPs/População(%)	Tamanho médio do bloco(KB)
MG	Belo Horizonte	2.375.151	14.012	0,58	620
	Uberlândia	654.681	3.990	0,61	482
	Juiz de Fora	550.710	2.272	0,41	522
	Montes Claros	390.212	575	0,15	348
RJ	Rio de Janeiro	6.320.446	29.049	0,46	663
	Petrópolis	298.017	403	0,14	500
	Volta Redonda	262.259	214	0,08	545
	Macaé	229.624	497	0,22	476
SP	São Paulo	11.895.893	54.256	0,46	585
	Campinas	1.154.617	6.586	0,57	598
	Ribeirão Preto	658.059	3.048	0,46	524
	Sorocaba	637.187	2.116	0,33	613

maiores populações sejam aquelas com mais clientes, não encontramos proporção clara entre o número de habitantes e a quantidade de espectadores.

## 5. Trabalhos Relacionados

Com o crescimento da Internet, também observamos um crescimento em sua complexidade. Eventos antes bem compreendidos podem não corresponder à realidade atual, face a escala que estes podem tomar. Por exemplo, há menos de 10 anos, grandes eventos de mídia ao vivo alcançavam cerca de 3,5 milhões de requisições de usuários em um período de quatro semanas [Veloso et al. 2002]. Hoje em dia, uma única transmissão, como a final de um campeonato de futebol, ou o discurso de um presidente, pode atrair a atenção de dezenas de milhões de usuários, em menos de uma hora<sup>5</sup>.

Um dos primeiros trabalhos a caracterizar transmissões de vídeo ao vivo na Internet foca no comportamento dos usuários que assistiam as transmissões [Veloso et al. 2002]. Especificamente, os autores analisaram a transmissão de um canal ao vivo na Internet e modelaram características dos seus usuários, como por exemplo, o tempo que eles permaneciam assistindo a transmissão ao vivo. Diferentemente daquele trabalho, em nosso trabalho estamos interessados em características do tráfego e seu impacto.

Grande parte das soluções para transmissão ao vivo na Internet utilizam distribuição em redes com arquitetura descentralizada (i.e., redes Par-a-Par, ou P2P) para melhorar sua escalabilidade. Por consequência, há uma série de trabalhos que avaliam tais sistemas e seus usuários [Hei et al. 2007, Jiangchuan Liu and Zhang 2008, Marfia et al. 2007, Chen et al. 2013]. Nessa linha, os autores focam no comportamento dos usuários e no desempenho da rede. Alguns desses trabalhos mostram que existem eventos de grande interesse que foram transmitidos em arquitetura P2P. Em alguns casos, os sistemas atenderam quase 200 mil usuários simultâneos [Hei et al. 2007]. O sistema que avaliamos não utiliza distribuição P2P, o que aumenta os problemas de escalabilidade e motiva estudos de caracterização de carga para o adequado provisionamento de capacidade.

Mais similar ao nosso trabalho, [Erman and Ramakrishnan 2013] fazem uma caracterização da transmissão ao vivo do Superbowl, o maior evento esportivo dos Estados Unidos. Porém, ao contrário do nosso trabalho, eles avaliam apenas a parcela do tráfego que passa por uma rede de telefonia celular.

Assim, apesar de existir uma série de trabalhos que estudam, caracterizam e mo-

<sup>5</sup><http://gigaom.com/2009/02/07/cnn-inauguration-p2p-stream-a-success-despite-backlash/>

delam alguns aspectos de transmissão de vídeo ao vivo na Internet, poucos focam no impacto gerado por um grande evento. Mais ainda, a maioria dos trabalhos é voltada para a análise do comportamento dos usuários. Neste trabalho nós focamos na carga imposta aos servidores e na qualidade de serviço aos clientes. Outros trabalhos enfocam a caracterização e modelagem do comportamento de usuários em sistemas de transmissão sob demanda de vídeo pré-armazenado. Por exemplo, alguns estudos abordaram propriedades do comportamento interativo destes usuários [Costa et al. 2004] e [Rocha et al. 2005]. Os autores avaliam os perfis de uso, com foco na interatividade do sistema por parte dos usuários, caracterizando por exemplo, o tempo de permanência e o padrão de execução das mídias. Transmissões ao vivo, foco do trabalho atual, têm características inerentemente diferentes de transmissões sob demanda. Logo, consideramos os esforços nesta direção complementares ao nosso.

## 6. Conclusões e trabalhos futuros

Transmissão de vídeo ao vivo na Internet é cada vez mais comum. Grandes eventos são difundidos ao vivo e, potencialmente, atraem um grande número de espectadores. Esse novo horizonte, com transmissões que, de repente, atraem um número de clientes acima do esperado impõe desafios para provisionamento de capacidade de provedores de serviço na Internet. Neste trabalho apresentamos uma caracterização da carga de trabalho de servidores do Globo.com durante transmissões ao vivo dos jogos da Copa das Confederações 2013. Nós avaliamos a arquitetura de serviços da Globo.com que, em alguns jogos, suporta carga muito elevada. Por exemplo, as transmissões avaliadas atendem até 180 mil sessões simultâneas e disseminam até 115 TB de dados em um único jogo com taxas de transmissão instantâneas que chegam a 350 GB/s.

Além disso, o número de espectadores, e consequentemente a carga imposta aos servidores, varia significativamente (1600%) entre os jogos com maior e menor carga. Isso ocorre, principalmente, em função da popularidade dos times e do horário da transmissão. Essa variação de carga pode levar a problemas de provisionamento. Por exemplo, mostramos que houve uma limitação na qualidade da mídia para adequar o consumo de banda e o número inesperado de clientes aos recursos existentes.

Caracterizamos também a qualidade da mídia recebida pelos usuários em função de seus provedores de acesso à Internet. Observamos que, grandes provedores de Internet banda larga no Brasil direcionam grande parte de seus clientes para o mesmo ponto de distribuição do Globo.com. Espectadores assinantes do maior provedor de acesso à Internet avaliado foram os que receberam mídia na menor qualidade. Também avaliamos a qualidade da mídia recebida pelos usuários em função de sua localização geográfica e IDH do estado. Nossos resultados mostram que a qualidade da mídia e a fração da população assistindo ao jogo pela Internet são correlacionadas com o IDH.

Em suma, nossos resultados confirmam os desafios para o provisionamento de capacidade em servidores de vídeo durante grandes eventos. Em trabalhos futuros pretendemos estudar técnicas para prever a carga durante grandes eventos. Também pretendemos avaliar o impacto da transmissão de grandes eventos ao vivo na infra-estrutura de rede, correlacionando a carga nos servidores com medições de rota obtidas com *traceroute*. Por último, estamos em contato com o Globo.com para recuperar e avaliar a carga da Copa do Mundo de Futebol 2014, também realizada no Brasil.

## 7. Agradecimentos

Os autores agradecem o apoio do CNPq, CAPES e da FAPEMIG e também a ajuda dos times de tecnologia da globo.com.

## Referências

- Cesario, M. V. (2012). Uso de anycast para balanceamento de carga na globo.com. *Talks and Tutorial, SBRC 2012*. Disponível em: <http://pt.slideshare.net/marcuscesario/apresentacao-anycast-sbrc201205>.
- Chen, Y., Zhang, B., Liu, Y., and Zhu, W. (2013). Measurement and modeling of video watching time in a large-scale internet video-on-demand system. *Multimedia, IEEE Transactions*, 15(8):2087–2098.
- Costa, C. P., Cunha, I. S., Borges, A., Ramos, C. V., Rocha, M. M., Almeida, J. M., and Ribeiro-Neto, B. (2004). Analyzing client interactivity in streaming media. In *Proceedings of the 13th international conference on World Wide Web*, pages 534–543. ACM.
- Erman, J. and Ramakrishnan, K. (2013). Understanding the super-sized traffic of the super bowl. *IMC '13 Proceedings of the 2013 conference on Internet measurement conference*, pages 353–360.
- Hei, X., Liang, C., Liang, J., Liu, Y., and Ross, K. W. (2007). A measurement study of a large-scale p2p iptv system. *IEEE Transactions Multimedia*, pages 1672–1687.
- Jiangchuan Liu, Sanjay G. Rao, B. L. and Zhang, H. (2008). Opportunities and challenges of peer-to-peer internet video broadcast. *Proceedings of the IEEE*, pages 11–24.
- Katabi, D. and Wroclawski, J. (2000). A Framework for Scalable Global IP-anycast. In *Proc. ACM SIGCOMM*.
- Mahanti, A. (2014). The evolving streaming media landscape. *Internet Computing, IEEE*, 18(1):4–6.
- Marfia, G., Sentivelli, A., Tewari, S., Gerla, M., and Kleinrock, L. (2007). Will IPTV ride the peer-to-peer stream? In *Communications Magazine, Special Issue on Peer-to-Peer Streaming*. IEEE.
- Rocha, M., Maia, M., Cunha, Í., Almeida, J., and Campos, S. (2005). Scalable media streaming to interactive users. In *Proceedings of the 13th annual ACM international conference on Multimedia*, pages 966–975. ACM.
- Shafiq, M. Z., Ji, L., Liu, A. X., Pang, J., Venkataraman, S., and Wang, J. (2013). A first look at cellular network performance during crowded events. *Proceedings of the ACM SIGMETRICS/international conference on Measurement and modeling of computer systems*, pages 17–28.
- Stockhammer, T. (2011). Dynamic adaptive streaming over http – standards and design principles. *MMSys '11 Proceedings of the second annual ACM conference on Multimedia systems*, pages 133–144.
- Veloso, E., Almeida, V., Meira, W., Bestavros, A., and Jin, S. (2002). A hierarchical characterization of a live streaming media workload. In *Proceedings of the 2nd ACM SIGCOMM Workshop on Internet measurment*, pages 117–130. ACM.