A Detailed Classification of Flash Events: Client, Server and Network Characteristics

Rashpinder Pal
Lecturer, Department of CSE
G.T.B.K.I.E.T. Chhapian Wali, Malout, Punjab, India
er.rashpinder@gmail.com

Sunil Kumar
Computer Faculty
Punjab ICT Society, Mansa, Punjab, India
sunil.budhlada@gmail.com

Rupinder Lali Sharma
Lecturer, Department of CSE
G.Z.S. College of Engg. & Tech., Bathinda, Punjab, India
rupinderlali89@gmail.com

Abstract— **Flash Events is the situation, when many users simultaneously access a computer server. Distributed denial-of-service (DDoS) attacks are similar to flash Events, but the only difference is users' intention. The attackers mimic Flash crowds to avoid their detection, thus it has been a rapidly growing problem in Network Security to accurately detect the DDoS attacks. Due to the overwhelming variety of DDoS attacks and Flash Events (FE), it is very difficult to distinguish them from one another. Our problem space of distinguishing Flash Events from DDoS has become large, but the solution space is yet very summarized, because very less work has been done on identifying the variety of Flash Events. This paper presents classification of Flash Events, and thus provides researchers with a better understanding of the problem space, so that it'd be easy for the researchers to work on distinguishing variety of DDoS attacks from variety of Flash Events.**

Keywords— DDoS, Flash Event, Classification, Distinguish, Network Security.

I. INTRODUCTION

The web servers experience high load during Flash event e.g., the World Cup web server during sale of tickets for World Cup Final, a university server at the time of results etc. The most important characteristic of Flash Event (FE) is that it is created by legitimate users; Whereas Denial of Service attacks attempt to make a Service unavailable to its intended users. These attacks are performed by a set of compromised machines called bots or zombies or slaves. The attacks in DDOS Scenario become geographically distributed and come from multiple sources at the same time thus are even more dangerous [1].

DDoS attacks try to slow down or block the server by consuming its CPU, Bandwidth or Memory resources. [2] Attackers are innovating powerful DDoS attack mechanism to avoid detection. [3] Moreover, attack tools are evolving and attackers mimic Flash crowds to avoid their detection. The attackers' use genuine IP addresses of the compromised machines so that the attack packets may look very similar to the ones sent by legitimate users, thus it becomes even harder by the victim server to detect the DDoS attacks [4]. DDoS attacks are nearly impossible to trace due to the large number of attack paths.

At the present time, the FE and DDoS attacks are very similar, and it is very difficult to distinguish them from one another, Both FE and DDoS are caused by a large number of requests which results in slow responses and connection drops [5]. It would be very beneficial for servers to distinguish between FE and DDoS attacks, because server deals both the events in a different way. During a FE the Content Distribution Networks and Locality of Reference mechanisms etc may be enabled to handle large number of users. In the case of a DDoS attack, the defense mechanisms are applied to block the attack traffic so that legitimate requests may be properly handled by the server.

In this paper we have presented a detailed taxonomy of Flash Events that may be helpful for the researchers to distinguish FE from DDoS attacks. This classification of Flash Events tries to answer many questions:

- How likely a Flash Event can occur?
- How much Load a Flash Event can pose on a Server?
- What are the traffic patterns and traffic behavior during a Flash Event?
- How much a FE may be distributed geographically?
- What is the behavior of users that cause FE?

Although this classification may not be very much detailed, yet the presented classification is nearly complete, because it covers almost all aspects, by which a Flash Event can be defined. The specified classification may be presented in other ways too i.e. the criteria and the classification techniques can be changed by changing the parameters based on requirements. This taxonomy does not propose any mechanism to distinguish between FE and DDoS attacks, so this paper is to be considered as a baseline to experiment on different types of Data Sets.

Section 2 provides information about related work in similar direction. Section 3 presents detailed classification

978-0-7695-4719-0/12 $26.00 © 2012 IEEE
DOI 10.1109/CSSS.2012.243

of Flash Events; Section 4 concludes the paper and Section 5 gives future work directions.

II. Related work

There were many research attempts in the past to classify Flash Events and to distinguish them from DDoS attacks; however those attempts were not particularly concentrated on classifying the Flash Events in much detail, also there has been little published work on categorization & classification of Flash Events.

The workload analysis of study of the 1998 Olympic Games Web site has been done in [6] that have been used to predict Flash Events at a Web site during its normal operation.

The Research work of Pan et al. shows characteristics of flash crowds, describes the design idea of FCAN and overviews its implementation issues [7]. The characteristics of Flash crowds require more elaboration and the same has been provided in this paper.

The World Cup Web site study presents a Flash Event analysis pointing to file referencing characteristic [8]. Also many clients repeatedly visited the site within short intervals. The web server's workload and their performance implications during Flash Events have been studied by Arlitt et al. [9].

Chen et al. [10] focused on detecting flash crowds caused by unpredictable events, which is a considerable and helpful measure towards describing the taxonomy of Flash Events.

Some research attempts focused on Web workload characterization which aims at understanding typical traffic behavior by studying traffic patterns observed by a web server. The SURGE [11] tool was developed to generate Web workload matching empirical measurements of several aspects of Web server usage.

Jung et al. [4] discussed the characteristic of Flash Crowds that the clients involved in Flash Events usually have browser history i.e. those clients visit other websites before Flash Events, however this work is concentrated towards CDNs rather than exploring Flash crowds characteristics.

Wendell et al.[12] characterized and quantified the behavior of thousands of flash crowds on CoralCDN, an open content distribution network running at several hundred POPs. Yet their characterization is restricted due to non consideration of several possibilities about real world Flash Crowds.

A study examined IP-flow-level traffic at the border router of the university network and presented flow characteristics of associated DDoS attacks and flash crowd behavior [13]. They found that flash crowd behavior was identified by quick increase in traffic to a particular flow followed by a step by step drop of packets over time. On the other hand DDoS attack had very sharp increase and decrease of the outgoing flows to a target server because of IP Spoofing, which appeared as a separate flow in the analysis, however this study did not targeted the areas where Flash Crowds may have similar characteristics to DDoS attacks.

Requests from spiders are another type of HTTP requests that may inconvenience a Web site and, while usually legitimate, might be genuine users for being dropped by an overloaded site.

III. Classification of Flash Events

We need to find out behavior of Flash Events after understanding their properties. To devise the detailed classification of Flash Events, we figure out nearly all the characteristics i.e. Flash Events' Predictability characteristics, Server Load characteristics, File Reference characteristics, Clients' characteristics and Traffic characteristics during FE. In this section we show this classification diagrammatically in Figure 1 and further we discuss and explain this classification. We characterize Flash Event with the following dimensions:

A. Classification of Flash Events by Client Characteristics

The clients accessing the web show up different behavior for different Web-Sites at different a time, which plays an important role in identifying the Flash Events. We can look for the following characteristics of a client involved in a Flash Event.

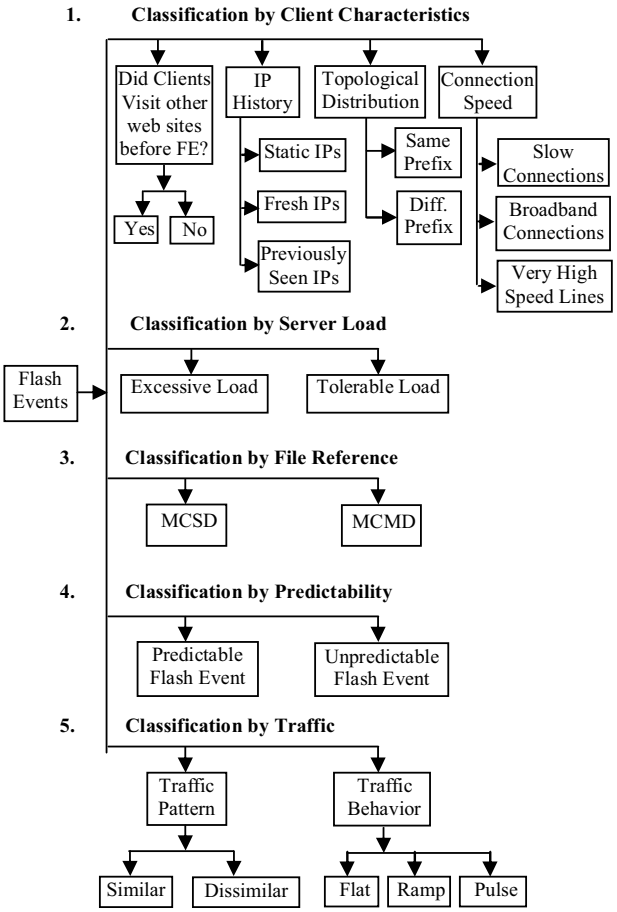
1) Did Clients Visit Other Web-Sites before FE (Browser History): During a Flash Event, it is a noticeable Point that the clients involved in Flash Event directly visit the victim website and their browser history may be Null, in such cases DDoS defenses may wrongly consider those Flash Events as Attacks, which may result in large user dissatisfaction, Such type of clients involved in flash events must be considered as Genuine, so that the server may take defensive measures accordingly.

2) IP History: Depending upon the No. of IPs and their Activation history, visiting history, three categories of Flash Events can be seen. The Flash Events Consisting of Large no. of Fresh IPs (The IPs that Visit the Web Site for the First time), such IPs are sometime discarded from accessing the web-page during FE due to trust issues, so our characterization highlights this category to be dealt with more accuracy. The other two categories of IPs (previously Seen IPs and Static IPs are trusted ones) are most common in Flash Events and these IPs are dealt with special mechanisms such as Load sharing algorithm, Traffic Shaping etc.

3) Connection Speed: The speed of requests being sent by a client plays an important role in determining the genuineness of the clients. Usually the requests arriving at Server, during a Flash Event come from all types of connections i.e. Slow Speed, High Speed and very high speed connections. But as far as the DDoS attacks are concerned, the attackers' connections have the similar speeds. Thus this characteristic helps in distinguishing DDoS from FE in a much better way.

4) Topological distribution of clients: Network-aware clustering technique [14] is used to determine the topological distribution of clients in Flash Events. Client

clustering allows us to collect individual clients into groups belonging to the same administrative domain. Clustering uses a large collection of unique network prefixes assembled from a wide set of BGP routing tables. The various clients' IP addresses are grouped into clusters based on longest prefix matching [4].



MCSD → Multiple Clients Access Single Document
MCMD → Multiple Clients Access Multiple Documents

Figure 1: Classification of Flash Events

B. Classification of Flash Events by Server Load

Generally it is believed that the load on a web server during a Flash Event is tolerable by means of traffic handling and congestion control mechanisms. The tolerable load does not necessarily mean normal condition, but a Flash Event that can be survived after applying traffic handling mechanisms or increasing network resources.

If the server load exceeds its maximum tolerance level, the server begins to slow down and can shutdown, such load is known as excessive load, which is very similar to DDoS attacks, and in such cases the Flash Events should be dealt carefully in order to avoid any wrong interpretation of Server Load. This wrong interpretation can cause denial of service to the legitimate clients due to preventive measures taken by server.

C. Classification of Flash Events by File Reference Characteristics

All the clients accessing the web have a common property that they request for a web page or a file, which can be generally called a document. The multiplicity of the clients is affirmed in Flash Events, as FEs can't occur due to a single client. The website can adapt Locality of reference mechanism, which enables reduction of server load through caching. Depending upon the document references made the FEs can be categorized as follows:

1) Multiple Clients Access Single Document (MCSD): It is well known that the clients involved in Flash Event generally access a single document [5] (e.g. Result page of a university website or a Popular TV Program). For Such type of Flash Events a Web site can provision some extra routes or mirrors for the destination document.

2) Multiple Clients Access Multiple Documents (MCMD): It is one of the areas that are neglected to be considered in Flash Events. During a Flash Event, clients may try to explore many areas of a web site, depending upon their requirements (e.g. Users may look for different applications of a Newly Launched operating System). Thus a web site experiencing Flash Event in such conditions may provision mechanisms to handle them better.

D. Classification of Flash Events by Traffic Characteristics

Traffic characteristics of the Web site are important because they determine provision of resources by server, to keep the site operational. Thus, watching traffic allows us to clearly express the period when a site can get overloaded and when the server can take defensive measures. The traffic characteristics can be categorized as Traffic Patterns and Traffic Behavior.

1) Traffic Patterns: The clients involved in Flash Events May have similar or dissimilar traffic patterns such as Packet type, Packet size, entropy, traffic flows etc. These patterns are very helpful in distinguishing FEs from DDoS attacks. They also help the server in provisioning the resources and applying defenses to cope with FEs or DDoS [4].

2) Traffic Behaviour: The traffic behaviour of Flash Event is usually Ramp or Pulse type, but in a few cases the traffic may show up Flat behaviour for some time, if the load is very high (e.g. Availability of World Cup Final tickets for few minutes, with prior news.) Such FEs can lead the server to shut down, for which the web site must provide the mechanism in advance to comply with such conditions.

E. Classification of Flash Events by Predictability

A flash event can be predictable [2] when a site is aware of the possibility of its occurrence. The Web site experiences a significant surge during the predicted time. (E.g. the on-line play-along Web site for a popular

television program.) Web site can provision in advance for such flash events and handle them better.

The unpredictable flash events [2] often arise due to the sudden load on a Web site unexpectedly. (E.g. Medical Web sites may experience unpredictable Flash Event during an epidemic or natural disaster, Popularity of a Web site due to failure of other similar web sites or News channels' web servers may get overloaded after some thrilling news).

Web sites often fail to correctly predict the demand of resources, even during predictable Flash Event. Both predictable and unpredictable flash events can thus pose a serious threat to a Web site.

IV. CONCLUSION

The DDoS field contains a variety of attacks, Flash Events and defense mechanisms, which hides certain issues regarding detecting and differentiating DDoS from FEs. This paper is an attempt to cut through the obscurity and structure the knowledge in this field. The proposed classification is intended to help the researchers think about the problems we face and the possible solutions. The proposed classification is by no means complete and all-encompassing. Some more issues may arise in future, some of which we cannot yet imagine. They will highlight new features for classification. Innovative approaches to DDoS detection and defense will be designed. They will also offer new design features carrying their share of benefits and weaknesses. We expect this classification to offer a foundation for classifying Flash Events and distinguishing them from DDoS attacks.

V. FUTURE SCOPE

While making the above classification, we selected those features of Flash Events and attack mechanisms that may offer critical information regarding correct identification of Flash Events so that the defense mechanisms may ensure that, no good users are treated as attackers. This classification will be helpful due to the following reasons:

A. A complete detail of Flash Events

For new researchers, this classification offers a comprehensive overview for a quick introduction to the Flash Event. Experienced researchers can use and extend this classification to structure and organize their knowledge in the field. This should lead to identification of new directions for research in distinguishing DDoS Attacks from Flash Events and improve understanding of the threat.

B. Improving the Existing Defense Systems

In addition to known differences and similarities between FEs and DDoS attacks, this classification explores a few situations seen rarely. As some FEs may wrongly be misunderstood as attacks by defense systems, this classification will help the researchers to improve defense systems.

C. Setting new benchmarks

The Flash Event classification will help the completeness of benchmark generation for DDoS defenses and DDoS/FE distinguishers. This classification will help expose and identify common weaknesses of a class of DDoS / FE distinguishers and DDoS defenses and design tailored experiments to remove these weaknesses.

D. Identifying unexplored research areas

Highlighting the effectiveness of different FE and DDoS distinguishers, defense mechanisms to handle attacks / Flash Events shows unexplored areas for research.

E. Cooperation among researchers

This classification will help the researchers in having a common view and similar ideas while designing DDoS attack detectors and defenses. It will also help in facilitating communication and offering a common language for discussing solutions. There is a need for the research community to develop common metrics and benchmarks for correctly detecting DDoS attack and properly evaluating DDoS defense.

RFERENCES

[1] C. Papadopoulos, R. Lindell, J. Mehringer, A. Hussain, R. Govindan, "COSSACK: Coordinated Suppression of Simultaneous Attacks".

[2] Bhatia, Sajal, Mohay, George M., Tickle, Alan, & Ahmed, Ejaz, " Parametric differences between a real-world distributed denial-of-service attack and a flash event."

[3] Kevin J. Houle, George M. Weaver, Neil Long & Rob Thomas, "Trends in Denial of Service Attack Technology"

[4] J. Jung, B. Krishnamurthy, and M. Rabinovich, "Flash crowds and denial of service attacks: Characterization and implications for CDNs and web sites," in Proceedings of the 11th international conference on World Wide Web, pp. 293–304, ACM, 2002.

[5] H. Park, P. Li, D. Gao, H. Lee, and R. Deng, "Distinguishing between FE and DDoS Using Randomness Check," Information Security, pp. 131–145, 2008.

[6] Arun K. Iyengar and Mark S. Squillante and Li Zhang. Analysis and Characterization of Large-Scale Web Server Access Patterns and Performance. World Wide Web, June 1999.

[7] Chenyu Pan, Merdan Atajanov, Mohammad Belayet Hossain, Toshihiko Shimokawa, and Norihiko Yoshida, "FCAN: Flash Crowds Alleviation Network," Proceedings of ACM 21st Annual Symposium on Applied Computing, pp.759-765, (April, 2006)

[8] Martin Arlitt and Tai Jin. Workload Characterization of the 1998 World Cup Web Site. HPL-1999-35R1

[9] Martin F. Arlitt and Carey L. Williamson. Internet Web servers: workload characterization and performance implications. IEEE\ACM Transactions on Networking, 5 (5):631--645, 1997.

[10] X. Chen and J. Heidemann. Flash Crowd Mitigation via Adaptive Admission Control based on Application-leve l Observations. USC/ISI Technical Report ISI-TR-557 (May 2002).

[11] Paul Barford and Mark Crovella. Generating Representative Web Workloads for Network and Server Performance Evaluation. In Measurement and Modeling of Computer Systems, 1998.

[12] Patrick Wendell, Michael J. Freedman. Going viral: flash crowds in an open CDN. In Proceedings of the ACM SIGCOMM Internet Measurement Conference (November 2011), pp. 549-558.

[13] Paul Barford and David Plonka. Characteristics of Network Traffic Flow Anomalies. In Proceedings of the ACM SIGCOMM Internet Measurement Workshop, Nov. 2001.

[14] Balachander Krishnamurthy and Jia Wang. On Network-Aware Clustering of Web Clients. In Proceedings of the ACM SIGCOMM, Aug. 2000.