

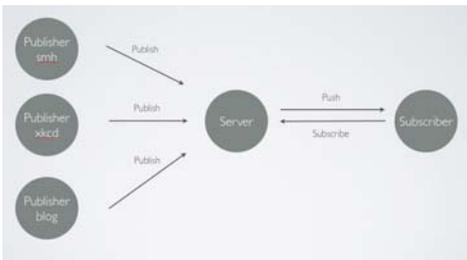
What is the problem?

Contextual Aware Systems

Pub-sub systems come in two basic forms:

- Content Based
- Topic Based

Topic-based pub-sub systems is a statically schemed system that offers limited expressiveness where subscribers subscribe to a particular node in the system. Content-based pub-sub systems on the other hand offer greater expressiveness as events are classified by its own properties.



Most pub-sub systems are topic based thus do not offer very much flexibility in determining how events are filtered.

The figure above outline the main interactions that a pub-sub system performs.

Scalability

One major issue that needs to be addressed is scalability. We must ensure that the system is capable of handling multiple requests by multiple users at the same time. Due to operating system and software constraints, most modern day web servers have the limits to only handle only ten thousand concurrent connections. Modern day hardware is no longer the bottleneck in networking throughput, but rather the implementation of the web frameworks that underpin web technologies.

Content Filtering

Message filters that are used in current pub-sub systems are implemented via keyword based searches and regular expressions. This is suitable for the average user, however, this system would fail for advanced users that are seeking to filter semi-structured messages such as html. We aim to create a system that offers greater expressiveness through the injection of python functions.

AIMS

Goals

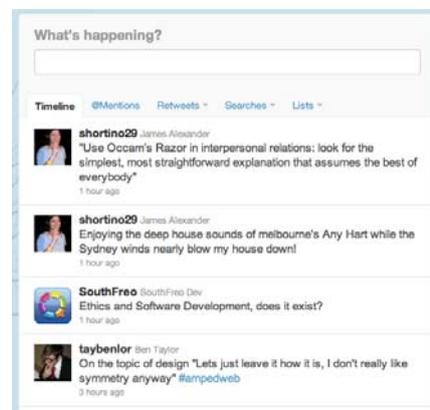
We aim to achieve the following goals in this research:

- Create a content-based system that is scalable and allows advanced filtering
- Create a novel filtering system based on injecting custom python functions

Motivation

Twitter

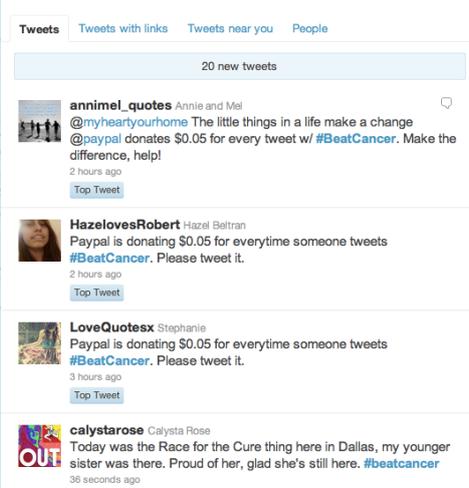
Twitter is a microblogging, social networking website that allows users to send messages to followers. Messages are defined as "tweets" and are in the form of 140 characters originally based on the limit of SMS text messages. It currently has over 160 million users and sends more than 90 million messages per day. Over the years, twitter has transformed from a simple status update mechanism to becoming a media, technical support and advertising platform, changing heavily the way the world communicates.



The image above shows a typical example of twitter's frontpage outlining the tweets posted by users.

We hope to provide more practical uses for pub-sub systems through the implementation of custom python functions during the filtering process.

Results for #beatcancer



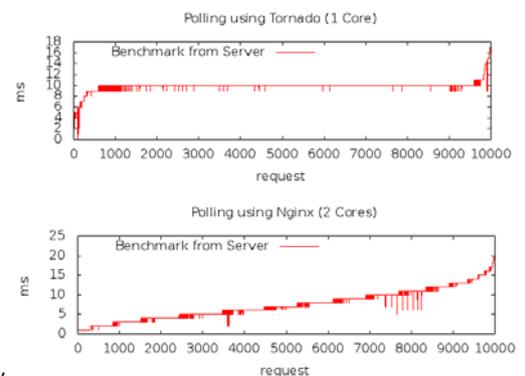
The image above outlines the use of #tags for filtering through messages in twitter.

Contributions

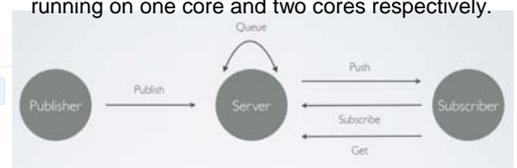
Contextual Aware Content-based Pub-Sub

By using pre-existing technologies and frameworks we will be developing an open source contextual based pub-sub system. Using this system, we can apply it to various applications such as managing stock prices, weather information or traffic data, but namely the Locator system for tracking people. Through the use of a sandbox environment, the custom filtering functions will be executed in a controlled environment safe from the rest of the system.

This is significant as we are providing an implementation of a pub-sub system that is not only used for message passing but can be adjusted for various other applications.



The diagrams above outline some basic performance tests conducted on our pub-sub system. 10,000 requests were made and comparisons were conducted on the system running on one core and two cores respectively.



The diagram above provides a top level outline of the basic methods that can be performed during its execution. Publishers and subscribers interact via a server and thus never directly interact with each other.<<

Future Work

We will continue to improve the system by focusing on areas that are parallelisable and finding faster ways of conducting event matching and filtering.