# Google Wave and Security

Jussi Litja
Helsinki University of Technology
jussi.litja@tkk.fi

## Abstract

Google Wave is a personal communication and collaboration tool based on a federated server architecture. This article overviews the Wave architecture and its features. Some data security aspects are considered as well. In addition to this, a comparison is made between Wave and some of the currently popular online communications services such as email, instant messaging and wikis. The comparison focuses on the benefits, drawbacks and security features of Wave compared to the other online services.

KEYWORDS: Google, Wave, security

## 1 Introduction

Internet has enabled the development of different online communication services such as email, instant messaging (IM), bulletin boards, newsgroups, wikis, Internet Relay Chat (IRC), etc. Most of these services are relatively old, but still remain very popular among Internet users. For example, the SMTP protocol used for sending email was introduced in the year 1982 [7]. However, on May 27, 2009 at the Google I/O conference, Google announced a new product that, it hoped, would merge these communication methods into a single service [3]. The new service is called Google Wave.

This article gives an overview of the Google Wave technology and the design principles behind it. First, the Wave architecture and some of its security features are introduced. Following this, a comparison is made between Wave and some of the currently popular online communication methods, such as email, instant messages, wikis, etc. The comparison focuses on the benefits and drawbacks of Wave compared to the other technologies, as well as on the security features of Wave compared to those found in the older services. Finally, some conclusions are made from the comparison results. But first, a short introduction to the Google Wave architecture and communication model.

## 2 Architecture

In most communication architectures, information is sent from one user to another in a unit called message. A message has typically a sender and one or more recipients. By contrast, in the Wave architecture users do not send messages. Instead, the system uses so called 'waves', which are, described by Google as, "equal parts conversation and document" [3]. A single wave can be thought of as a document that can be edited and commented concurrently by multiple users. Fig 1 shows the main differences between email and wave messaging from a high and abstract level. In email (on the left), each message has to be individually delivered to each of its receivers. For example, an email with 100 receivers results in 100 separate messages.

However, in Wave the communication principles are a bit different: waves are stored at 'wave providers' (Fig 1, on the right, between the users), rather than directly delivered from one user to another. That is, instead of receiving a message, the user simply receives a 'view' of the wave from his wave provider. There can be many wave providers, each of which is responsible for the users in its domain [10].

As an example of a wave provider, Google's own 'wave.google.com' is a wave provider available as a preview for a limited set of users. Still, despite the name 'Google Wave', anyone can become a wave provider. For example, a company can become a wave provider for its employees. In this sense, assigning domains to wave providers shares some resemblance to assigning mail address domains to email servers.

The wave providers communicate with each other to maintain copies of the waves - this is needed when a wave has participants from two (or more) different domains. In other words, when a wave is created or modified in the domain of some wave provider, the wave provider has to make sure that each participant of that wave is notified of those changes [10].

### 2.1 Waves and wavelets

Waves are defined as "hosted documents...supporting concurrent modifications and low-latency updates" [10]. A user can start a new wave and add participants to it, similarly as he would add recipients to an email. The participants receive a view of the wave and can start editing it - the modifications are updated to all of the other participants. In this sense a wave is similar to a shared document that can be edited concurrently.

However, apart from just editing the wave, the participants can add comments, or replies, to it. Comments can be shown to all of the participants or just to a selected subset of them. In the latter case the comments can be seen as 'private messages', similar to instant messages. Another interesting feature is that the modification history of a wave is stored persistently. This 'playback' feature can be used to observe how the wave has been modified over the course of time.

A wave consists of a set wavelets, as shown in Fig 2. Each of these wavelets has a list of participants (users) and a set of related documents that make up its contents [10]. When

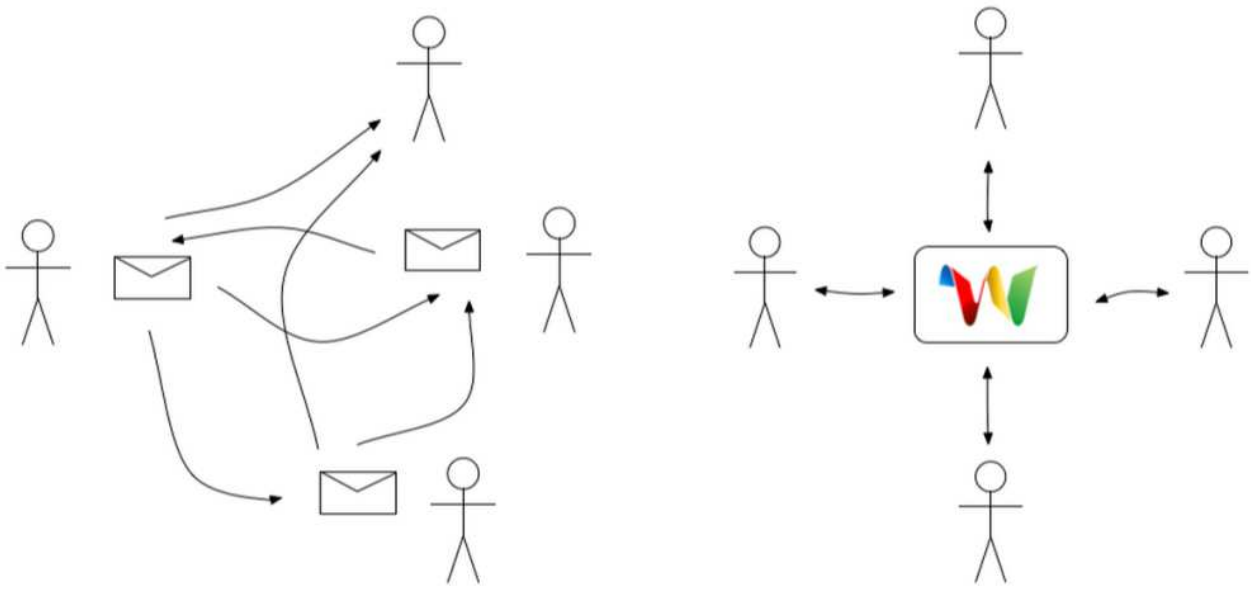

Figure 1: Email and Wave communication principles. [6]

a user opens a wave he receives a view of those wavelets (of that wave) that he is a participant of. This is quite different from, for example, reading email: an email message is fetched from a mail server to the user's computer and cannot be modified after it has been sent (or received).

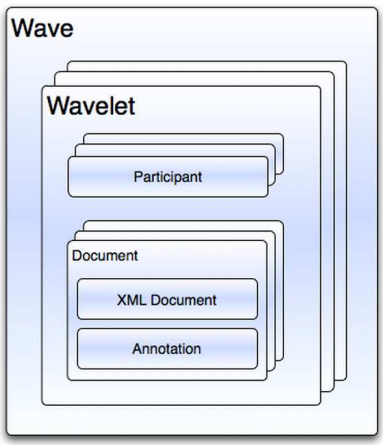


Figure 2: The anatomy of a wave (the Wave data model). [9]

A participant of a wave is not necessarily a participant of all the wavelets in that wave. This is the case, for example, when there is a private reply in the wave (that is not intended for the user).

## 2.2 Wave providers

As mentioned earlier, waves are stored at wave providers. In the wave architecture, each user has a wave address (e.g. username@domain, notice the resemblance to email addresses) which consists of a user name and domain name of a wave provider. When a user creates a wave it is stored at the wave provider responsible for that user's domain. However, if the wave has participants from other domains, a copy of the wave must be delivered to the wave providers responsible for the domains of the other participants. Similarly, when a wave is modified the changes must be propagated to all wave providers maintaining a copy of that wave. [10]

## 2.3 Wave service federation

A wavelet is called a 'local wavelet' when it is located at a wave provider where it originated from. Otherwise it is called a 'remote wavelet'. Now, if a user modifies a wave in its domain, how are the remote wavelets of that wave updated?

Wave service federation is an architecture where wave providers responsible for different domains enable users in those domains to communicate with each other, or in other words, provide an end-to-end communication service for all the users. The protocol used between the different wave providers is called the Google Wave Federation Protocol. The protocol specifies, among other things, how wave operations are forwarded from one wave provider to another and back. [10]

## 2.4 Robots and gadgets

Waves are shared documents that can be edited, viewed and commented concurrently. In addition to these basic features, there are extensions available that can be installed and used to broaden Wave's functionality: robots and gadgets.

A robot is an automated participant (user) on a wave. Robots can "modify information in a wave, interact with participants on a wave, communicate and synchronize information in a wave to the outside world or to other waves and access or modify state in a third-party system (such as a database)" [4]. Simply put, a robot can be thought of as a

program that is 'wave-aware', i.e. it knows the content and participants on a wave and can interact with them or take other actions based on them. An example of a robot in action is the spell-checking robot used by Google's own wave provider 'wave.google.com'. If a user creates a wave then the spell-checking robot "joins" it or in other words, becomes a participant in it. It then underlines the words that it does not recognize thus modifying the wave itself. Other examples of robots could be translation, chat and chess-playing robots. Moreover, robots can be developed for other, more customized purposes by using the Google Wave API [2].

The second type of extensions, gadgets, are small applications that can be embedded on waves. The difference between robots and gadgets is that while a robot is a participant on a wave, a gadget is simply an application and is contained, or lives in the wave [2]. A gadget cannot influence the wave in any way, but has an internal state that is shared with all the participants of the wave. As an example, a simple gadget could be a whiteboard application that is embedded to a wave and where the participants of that wave can draw on it. Or, gadgets could be small multi-player games such as a multi-player Sudoku game or utility gadgets such as polls (which the participants vote in). Like robots, gadgets can also be created by using the Google Wave API [5].

# 3 Security in Wave

The CIA triad is a common model used to describe different aspects of data security. In the CIA model the security of an information system can be evaluated by three factors: data confidentiality, integrity and availability. Confidentiality of data is a term meaning that the data cannot be disclosed to unauthorized users, i.e. the information cannot be viewed by unauthorized users. In turn, integrity means that the data cannot be modified by unauthorized users. Finally, availability refers to the fact that information needs to be available to the users - confidentiality and integrity are useless if the data is unavailable. This section overviews how data confidentiality, integrity and availability are considered in Wave and also presents few other information security aspects of Wave.

## 3.1 Confidentiality in Wave

The Google Wave Federation Protocol is an open extension to the XMPP Internet Messaging protocol [8]. Among a few other things adopted from XMPP to the federation protocol is authentication: wave providers communicating with each other must authenticate themselves using TLS [1]. However, it does not require that the traffic between them is encrypted, although it is recommended. Moreover, it is the responsibility of the wave provider to authenticate users in its domain and perform local access control. In conclusion, if the traffic between wave providers is encrypted and end-users are authenticated by wave providers then a reasonable level of confidentiality can be obtained. [10]

## 3.2 Integrity in Wave

The classic meaning of data integrity is that the data cannot be modified by unauthorized users. In Wave, if a wave has participants only in one domain, this can be achieved by performing local access control in the domain of the wave, i.e. authenticating the users. Then the wave cannot be modified by unauthorized users. However, if a wave has participants in multiple domain, then the wave operations must be transferred between wave providers. In that case the integrity of data can be guaranteed by using cryptographic signatures and certificates between the wave providers [10].

## 3.3 Availability in Wave

Availability is an important aspect of data security. Since using Wave requires an Internet connection, or at least a connection to the wave provider, this could very well be an issue with Wave. If there is no connection, the service is useless. However, the same applies also to other online communication methods, such as email and instant messaging - if there is no Internet connection they are not any more useful than Wave.

In Wave, if a participant of a wave is in a domain whose wave provider is unavailable, then the wave operations (modifications, etc.) cannot be sent to it. In this case the operations are queued so that when the wave provider does recover, they are sent to it.

Related to data availability is data persistence. Recording the version history of a wave (playback) may not always be a good thing. For example, when dealing with confidential information.

# 4 Comparison

The communication services compared here have all relatively different communication paradigms, i.e. they are used and meant to be used quite differently. However, some properties can be evaluated and compared. These include communication latency (speed), the feeling of presence, richness of the communication and a few other, service-specific features that are present in some of these communication services. In addition, the security aspects of the services can be evaluated against each other.

In this section I have compared Wave to few other popular online communication services, more specifically to email, instant messaging, wikis and IRC.

## 4.1 Wave vs. email

Email messages have a sender and one or multiple recipients. A separate message is sent to each recipient and once a message has been sent it cannot be modified. Probably the most notable difference between email and Wave is that Wave is based on the idea of served, hosted conversations whereas email is based on routing individual messages from sender to receiver.

The content of waves can be richer that email messages, since waves can be complemented with different extensions. On one hand, you may add attachments such as pictures and

videos to email messages. On the other hand, these may also be added to waves. Furthermore, waves can also contain gadgets that are small applications which allow interaction between the wave users in the wave. In this sense the content of waves can be considered to be richer than email messages.

Also, the feeling of presence is stronger when using waves, since the latency is smaller than it is with email messages. However, in some situations it may not be a preferable thing.

Considering security aspects, email users can be authenticated by the mail server they are using and the message transfer between then can be encrypted. However, when email messages are transferred between email servers, they are usually not encrypted, although it is possible. In this sense the security of email messaging is weaker than Wave messaging.

### 4.2 Wave vs. instant messaging

Instant messaging is a form of communication where end-users send each other messages either trough a proxy server of the service provider or directly via a peer-to-peer network. The nature of waves is quite different to IMs, since waves are hosted documents. Even though Wave allows comments to be written in waves, the conversation seem a bit forced and artificial, since 'instant messaging' in Wave is basically writing comments to a document, not truly taking part in a conversation.

IM users can be authenticated and IMs can be encrypted, but most of the popular IM services, such as AIM and MSN messenger do not encrypt messages. In this sense Wave is more secure than some of the instant messaging services.

### 4.3 Wave vs. wikis

Wikis are websites that contain interlinked web pages, i.e. the wiki pages contain links to other wiki pages. Wikis are typically used by communities and companies that share information internally and need to be able to modify that information. Waves contain some similarities to wikis: they are shared by group of people (participants) that can view and edit them. Waves can also be linked to other waves using so-called Wave IDs. Wikis often have a version history similar to Wave's playback option. These similarities show that some of Wave's functionality might have been inspired by wikis.

However, wikis lack the interactivity of waves. When a user types in a character to a wave it is shown to the other participants in near real-time. By contrast, in wikis the user often simply edits a single, 'static' version of the wiki page. If another user is editing the same page at the same time the resulting page is often a merged combination of these two pages. This merging may results in a conflict, especially when there are more than two users editing the same wiki page. A way to overcome this is to prohibit anyone from editing the page if somebody else is editing it, but this in turn causes unnecessary waiting for the other users. In conclusion, waves support letter-by-letter concurrent modifications, while editing a wiki page concurrently may result in unnecessary waiting or possible merge conflicts.

Since wikis are web pages, they can be secured with TLS or SSL, as can be waves.

### 4.4 Wave vs. IRC

IRC is an text-based online communication network, which consists of networked IRC servers and users connected to them. Users can join 'channels' in which they can write text messages to other participants in the channel. In this sense IRC is more interactive than say, email, and more like replies in wave. Nevertheless, IRC messaging is by nature "chat-like", meaning that once you write something to a channel it cannot be removed, in other words the flow of information is linear in time. This nature of messaging implies that IRC was not meant for collaboration tool, unlike wikis and for example, Wave.

Although some IRC networks support encrypted connections via SSL, the majority of them send the messages in plain-text. This means that when compared to Wave the security properties of IRC are much weaker that Wave's.

## 5 Conclusions

Google Wave is a new kind of approach to online communication. It is based on hosted conversations that can be edited and commented by multiple users. The modifications can be done concurrently and without notable latency problems. Waves are editable documents that can be commented and extended for example, with gadgets or robots. In that sense waves are a mixture of wikis, email and instant messaging.

The security features of Wave provide a reasonable level of security: the Google Wave Federation protocol enables authentication and encryption of messages delivered between wave providers. This combined with good local access control by wave providers results in fairly good end-to-end security between wave users from different domains.

Although the features of Google Wave seem exciting (at least to the writer) it is hard to know whether it will ever be adopted widely. I have no doubt that Wave would be an efficient collaboration tool for people that write documents together. For that purpose it's features are quite ideal. Moreover, scheduling events via email tends to build up a chain email consisting of dozens of replies. This kind of iterating could possibly be done more easily with waves. Then again, to use Wave as a instant messaging tool seems a bit artificial.

In conclusion, the Wave is an interesting new service that tries to bring forth a new kind of communication and collaboration paradigm. However, since Wave is currently available only as a limited preview it is difficult to know whether it will be adopted by the general population.

## References

[1] T. Dierks and E. Rescorla. The Transport Layer Security (TLS) Protocol Version 1.2. RFC 5246 (Proposed Standard), Aug. 2008. Updated by RFC 5746.

[2] Google. Wave extensions, 2009. `http://code.google.com/apis/wave/extensions`.

[3] Google. About google wave, 2010. http://wave.google.com/about.html.

[4] Google. Google wave robots api, 2010. http://code.google.com/apis/wave/extensions/robots/index.html.

[5] Google. Wave gadgets tutorial, 2010. http://code.google.com/apis/wave/extensions/gadgets/guide.html.

[6] D. Peterson. Google wave federation day info, 2009. http://www.waveprotocol.org/presentations.

[7] J. Postel. Simple Mail Transfer Protocol. RFC 821 (Standard), Aug. 1982. Obsoleted by RFC 2821.

[8] P. Saint-Andre. Extensible Messaging and Presence Protocol (XMPP): Core. RFC 3920 (Proposed Standard), Oct. 2004.

[9] Soren Lassen. Google wave federation architecture overview, 2009. http://www.waveprotocol.org/presentations.

[10] Soren Lassen, Sam Thorogood. Google wave federation architecture. Technical report, Google, 2009. http://www.waveprotocol.org/whitepapers/google-wave-architecture.