**[Git] Git và Github là gì? Tại sao nên dùng?**

Cập nhật lần cuối ngày 25/04/2015 trong [Công cụ](http://thachpham.com/category/tools), [Web Development](http://thachpham.com/category/web-development)

Bài này thuộc phần 2 của 11 phần trong serie [**Git cơ bản**](http://thachpham.com/series/git-co-ban)

Một câu hỏi mà nhiều người đặt ra khi nhìn thấy serie này đó là Git là cái gì? Nó là ngôn ngữ lập trình gì chăng? Hay một loại thức ăn nào đó? Một kỹ thuật thiết kế website sao? Hoàn toàn không, bạn sẽ hiểu được Git là cái gì sau khi đọc hết từng chữ trong bài viết này.

Trước khi nói qua về Git, mình có một câu chuyện ngắn muốn kể cho các bạn như sau.

Có hai anh chàng lập trình viên đang hợp tác với nhau cùng xây dựng một phần mềm cho khách hàng. Dĩ nhiên hai anh chàng này sẽ làm việc với nhau bằng cách mỗi người tự viết code một ít và gửi cho nhau qua email hoặc một phương thức gửi dữ liệu thông thường nào đó. Mỗi người sau khi nhận được code của nhau sẽ tiến hành tự xem và sửa lại, sau đó cùng nhau ráp nối vào phần mềm mà họ cần làm. Tuy nhiên, một hôm anh A bỗng nhiên vô tình viết code sai vào phần mềm dẫn tới sản phẩm bị lỗi mà trước đó cả hai anh đều không hề sao lưu lại do phần mềm quá lớn. Và thế là hai anh lại cùng nhau xây dựng lại tính năng bị lỗi và tiếp tục gửi cho nhau như vậy.

Thế bạn có thấy vấn đề gì xảy ra ở câu chuyện trên không? Mình thì mình thấy có hai vấn đề như sau:

1. Việc tự code riêng và gửi cho nhau qua email rất mất thời gian của nhau. Giá như anh A có thể chủ động xem những thay đổi của anh B từ xa và tiến hành gộp trực tiếp những thay đổi của anh B vào sản phẩm.
2. Việc sửa code mà không sao lưu khiến cho họ phải viết lại code từ đầu khi phần mềm bị lỗi.

Và hai vấn đề trên Git có thể sẽ giúp hai anh chàng ấy giải quyết nhanh gọn lẹ và đơn giản hơn rất nhiều. Ngoài ra Git còn làm nhiều việc hơn nữa mà chỉ có thể sử dụng bạn mới hiểu được.

**Git là gì?**

**Git** là tên gọi của một **Hệ thống quản lý phiên bản phân tán** (*Distributed Version Control System –* ***DVCS***) là một trong những hệ thống quản lý phiên bản phân tán phổ biến nhất hiện nay. DVCS nghĩa là hệ thống giúp mỗi máy tính có thể lưu trữ nhiều phiên bản khác nhau của một mã nguồn được nhân bản (**clone**) từ một kho chứa mã nguồn (**repository**), mỗi thay đổi vào mã nguồn trên máy tính sẽ có thể ủy thác (**commit**) rồi đưa lên máy chủ nơi đặt kho chứa chính. Và một máy tính khác (nếu họ có quyền truy cập) cũng có thể clone lại mã nguồn từ kho chứa hoặc clone lại một tập hợp các thay đổi mới nhất trên máy tính kia. Trong Git, thư mục làm việc trên máy tính gọi là **Working Tree**. Đại loại là như vậy.

[](http://s3cdn.thach.io/wp-content/uploads/2015/04/21191808/dvcs.png)

Mô hình hoạt động của DVCS

Ngoài ra, có một cách hiểu khác về Git đơn giản hơn đó là nó sẽ giúp bạn lưu lại các phiên bản của những lần thay đổi vào mã nguồn và có thể dễ dàng khôi phục lại dễ dàng mà không cần copy lại mã nguồn rồi cất vào đâu đó. Và một người khác có thể xem các thay đổi của bạn ở từng phiên bản,  họ cũng có thể đối chiếu các thay đổi của bạn rồi gộp phiên bản của bạn vào phiên bản của họ. Cuối cùng là tất cả có thể đưa các thay đổi vào mã nguồn của mình lên một kho chứa mã nguồn.

Cơ chế lưu trữ phiên bản của Git là nó sẽ tạo ra một “*ảnh chụp*” (***snapshot***) trên mỗi tập tin và thư mục sau khi commit, từ đó nó có thể cho phép bạn tái sử dụng lại một ảnh chụp nào đó mà bạn có thể hiểu đó là một phiên bản. Đây cũng chính là lợi thế của Git so với các DVCS khác khi nó không “lưu cứng” dữ liệu mà sẽ lưu với dạng snapshot.