

Đặt tả hệ thống

Hệ thống trang báo điện tử

- Có khoảng 10 tác giả (author) - đăng tải các bài báo (article) hằng ngày
- Có khoảng 10 M độc giả (user) – đọc bài báo và đưa bình luận (comment) của bài báo
- Có 10 nhà phân tích – chuyên phân tích các article, comment, lượng view bài báo

Yêu cầu hệ thống:

- Thu thập và lưu trữ được các dữ liệu về tác giả, độc giả, bài báo, bình luận, lượng view.
- Phân tích và thống kê ra được xu hướng quan tâm của độc giả, xu hướng các bài báo của các tác giả

Một số Kịch bản của bài toán và thống kê

- Khi tác giả đăng các bài báo thì thông tin về việc đăng tải bài báo được tự động gửi về server
- Lượng view của bài báo hoặc của độc giả được tự động gửi về hệ thống khi bài báo được độc giả click chọn xem.
- Xu hướng quan tâm của độc giả dựa vào lượng của loại hình bài báo mà độc giả thường xuyên đọc trong 1 giờ/ 1 ngày/1 tuần/1 tháng.
- Xu hướng các báo dựa vào loại hình bài báo được các tác giả thường xuyên đăng tải trong 1 giờ/1 ngày/1 tuần/1 tháng.
- Số lượng comment của các độc giả trên bài báo cũng thể hiện sự quan tâm của độc giả về bài báo đó.
- Một nhà phân tích có thể chạy 10 câu truy vấn mỗi giờ để lấy dữ liệu thông tin cần phân tích.
- Theo thống kê thì mỗi tác giả có thể đăng tải 1 bài báo trong 1 ngày; mỗi bài báo mỗi ngày có thể có đến 1000 lượt bình luận; lượng view các bài báo 10 M mỗi ngày ;
- Những thông tin được độc giả quan tâm thì lưu trữ trong 3 năm, tất cả các thông tin được lưu trữ 10 năm.
- Thông tin một bài báo cần được lưu trữ: tiêu đề, tóm tắt, ngày giờ đăng, lượng like, lượng view, URL, tác giả, loại. Thông tin bình luận cần được lưu trữ: nội dung bình luận, người bình luận, ngày giờ, lượng like.

Yêu cầu bài tập: Phân tích và đặt tả workload cho hệ thống trên (chương 2)

- (1) Đưa ra danh sách các CRUD của hệ thống, ứng với actor, Operation, Information, Frequency, criticality,
- (2) Phân tích chi tiết 2 thao tác (operation) nào mà quan trọng nhất.
- (3) Đưa ra các thực thể lưu trữ dữ liệu (tên thực thể, cấu trúc lưu trữ), mối quan hệ giữa chúng (mô hình khái niệm).
- (4) Đưa ra data size của document lưu trữ ứng với thao tác ghi (write) quan trọng.