

GIẢI PHÁP ĐẢM BẢO AN TOÀN DỮ LIỆU CHO CÔNG DỮ LIỆU MỞ TỈNH THANH HÓA

APPLICATION OF BLOCKCHAIN TECHNOLOGY FOR INFORMATION SECURITY IN OPEN DATA PORTALS

Trịnh Sỹ Tuấn^{1*}
Trịnh Viết Cường²

¹ Khoa Kỹ thuật Công nghệ, Trường Đại học Hồng Đức;

² Phòng Khoa học Công nghệ và Hợp tác Quốc tế, Trường Đại học Hồng Đức

*Tác giả liên hệ; Email: trinhtsyuan@gmail.com

DOI: <https://doi.org/10.70117/hdujs.3030-4628.90.09.2026.1244>

Ngày nộp bài: 01/04/2026; Ngày gửi phản biện: 06/04/2026; Ngày duyệt đăng: 25/09/2026

Từ khóa:

*Blockchain, Công
dữ liệu mở, An toàn
thông tin, Tính toàn
vẹn dữ liệu, Phi tập
trung hóa.*

Keywords:

*Blockchain, Open
Data Portal,
Information
Security, Data
Integrity,
Decentralization.*

TÓM TẮT

Chuyển đổi số đặt ra yêu cầu cấp thiết về an toàn, toàn vẹn dữ liệu cho các cổng dữ liệu mở trước nguy cơ tấn công mạng và sửa đổi trái phép. Nghiên cứu đề xuất, thử nghiệm mô hình tích hợp Blockchain (Hyperledger Fabric, giao thức Raft) vào kiến trúc Cổng dữ liệu mở tỉnh Thanh Hóa.

Mô hình đảm bảo toàn vẹn và truy vết dữ liệu; hệ thống đạt thời gian tạo khối 142-547 ms, độ trễ xác thực dưới 200 ms (tệp ≤10 MB), thông lượng ổn định đến 534 giao dịch/giây với 500 người dùng đồng thời.

Giải pháp nâng cao độ tin cậy, minh bạch và khả năng chống giả mạo cho cổng dữ liệu mở, có thể nhân rộng cho các hệ thống dữ liệu mở cấp tỉnh khác.

ABSTRACT

Digital transformation raises urgent requirements for data safety and integrity in open data portals amid cyber-attacks and unauthorized modification. This study proposes and tests a model integrating Blockchain (Hyperledger Fabric, Raft protocol) into the architecture of the Thanh Hoa Open Data Portal.

The model ensures data integrity and traceability; the system achieves a block-creation time of 142-547 ms, validation latency under 200 ms (files ≤10 MB), and stable throughput of up to 534 transactions/second with 500 concurrent users.

The solution improves reliability, transparency and tamper-resistance for open data portals, and can be extended to other provincial-level open data systems.

1. ĐẶT VẤN ĐỀ

1.1. Giới thiệu về cổng dữ liệu mở OpenData tỉnh Thanh Hóa

Dữ liệu mở (Open Data) là dữ liệu được công bố công khai, cho phép bất kỳ ai cũng có thể tự do sử dụng, tái sử dụng và phân phối mà không bị giới hạn bởi bản quyền, quyền sở hữu hay bất kỳ ràng buộc pháp lý nào khác. Loại dữ liệu này thường được phát hành dưới dạng tệp tin điện tử, có thể dễ dàng truy cập và tải xuống từ các nền tảng trực tuyến. Nguồn gốc của dữ liệu mở rất đa dạng, từ các cơ quan chính phủ, tổ chức phi chính phủ, đơn vị phi lợi nhuận cho đến cộng đồng người dùng. Điểm cốt lõi của khái niệm này là dữ liệu phải được công bố theo các tiêu chuẩn mở, đảm bảo mọi người đều có thể tự do khai thác, phân tích và tái sử dụng vì bất kỳ mục đích nào. Dữ liệu mở được hình thành và phát triển mạnh mẽ trong những năm gần đây, đặc biệt trong bối cảnh chuyển đổi số và xây dựng chính phủ điện tử. Ngày nay, dữ liệu mở đóng vai trò quan trọng trong việc bảo đảm tính minh bạch, thúc đẩy đổi mới sáng tạo và nâng cao hiệu quả quản lý nhà nước. Trong dữ liệu mở, các cơ quan, tổ chức và đơn vị công khai các bộ dữ liệu thuộc nhiều lĩnh vực khác nhau như kinh tế, giáo dục, y tế, môi trường hay giao thông để bất kỳ ai đều có thể truy cập và khai thác. Tuy nhiên, nếu dữ liệu chỉ được công bố dưới dạng rời rạc, thiếu tiêu chuẩn hoặc không rõ nguồn gốc thì người sử dụng sẽ gặp khó khăn trong việc kiểm chứng độ tin cậy và tính chính xác. Điều này đặt ra yêu cầu cần có một cơ chế quản lý và công bố dữ liệu theo các tiêu chuẩn mở, bảo đảm dữ liệu có cấu trúc rõ ràng, định dạng máy có thể đọc được và cho phép tái sử dụng tự do. Cổng dữ liệu mở là nền tảng trực tuyến trung tâm, được xây dựng nhằm phục vụ nhu cầu tiếp cận, khai thác và tái sử dụng dữ liệu của người dân, doanh nghiệp và các tổ chức. Cổng dữ liệu mở tỉnh Thanh Hóa có các chức năng như sau:

Công bố và lưu trữ dữ liệu: Đây là chức năng cốt lõi của cổng. Các cơ quan, đơn vị nhà nước có thể đăng tải, cập nhật và quản lý các bộ dữ liệu thuộc phạm vi quản lý của mình. Mỗi bộ dữ liệu được lưu trữ kèm theo thông tin mô tả đầy đủ như tên, đơn vị cung cấp, lĩnh vực, định dạng tệp và giấy phép sử dụng, giúp người dùng nắm rõ nguồn gốc và phạm vi áp dụng của dữ liệu trước khi tải về.

Phân loại và tổ chức dữ liệu: Cổng tổ chức dữ liệu theo nhiều chiều phân loại khác nhau, gồm theo đơn vị cung cấp, theo lĩnh vực (giáo dục, kinh tế, y tế, văn hóa - du lịch, nông nghiệp, giao thông vận tải, tài nguyên - môi trường...), theo từ khóa và theo định dạng tệp. Cách tổ chức này giúp người dùng dễ dàng định hướng và tìm đến đúng nguồn dữ liệu cần thiết.

Tìm kiếm và tra cứu dữ liệu: Cổng cung cấp công cụ tìm kiếm cho phép lọc và tra cứu dữ liệu theo nhiều tiêu chí như đơn vị, lĩnh vực, khoảng thời gian và định dạng tệp. Ngoài ra, chức năng dữ liệu tra cứu cho phép người dùng truy vấn trực tiếp các danh mục dữ liệu có cấu trúc mà không cần tải tệp về, đặc biệt hữu ích với các bộ dữ liệu lớn hoặc được cập nhật thường xuyên.

Gửi và đóng góp dữ liệu: Cổng hỗ trợ chức năng gửi dữ liệu lên hệ thống, cho phép các đơn vị và tổ chức chủ động đóng góp vào kho dữ liệu chung. Điều này góp phần mở rộng nguồn dữ liệu, tăng tính đa dạng và phong phú của hệ sinh thái dữ liệu mở trên địa bàn.

Thông tin và truyền thông: Bên cạnh dữ liệu, công còn tích hợp chuyên mục tin tức và tin tổng hợp, cập nhật các thông tin liên quan đến chuyên đổi số, chính sách dữ liệu mở và các hoạt động triển khai của tỉnh. Đây là kênh truyền thông hữu ích giúp người dùng nắm bắt định hướng phát triển và các quyết định mới nhất liên quan đến dữ liệu mở.

Nhờ tích hợp đầy đủ các chức năng trên, cổng dữ liệu mở không chỉ là kho lưu trữ thụ động mà còn là một nền tảng chủ động, góp phần thúc đẩy văn hóa chia sẻ dữ liệu, nâng cao tính minh bạch trong quản lý nhà nước và tạo điều kiện cho người dân, doanh nghiệp khai thác dữ liệu phục vụ các mục tiêu phát triển kinh tế - xã hội.

1.2. Tại sao cần đảm bảo an toàn dữ liệu cho Cổng dữ liệu mở

Hiện nay, hệ thống cổng dữ liệu mở tỉnh Thanh Hóa đang sử dụng kiến trúc Cơ sở dữ liệu truyền thông (cơ sở dữ liệu tập trung). Với kiến trúc này có hai nhược điểm lớn về an toàn dữ liệu:

Tính toàn vẹn của dữ liệu: dữ liệu trong hệ thống cơ sở dữ liệu tập trung dễ dàng bị chỉnh sửa bởi nhà quản trị cơ sở dữ liệu, cũng như kẻ tấn công bên ngoài. Do đó dữ liệu có thể bị chỉnh sửa trái phép so với mong muốn của người cung cấp dữ liệu lên Cổng dữ liệu. Việc chỉnh sửa dữ liệu trái phép sẽ có thể dẫn tới những người sử dụng dữ liệu trên Cổng dữ liệu mở như người dân và doanh nghiệp có thể đưa ra quyết định sai lầm, gây hậu quả nghiêm trọng.

Tính sẵn sàng của dữ liệu: hệ thống cổng dữ liệu mở tỉnh Thanh Hóa đang sử dụng kiến trúc cơ sở dữ liệu tập trung (mô hình khách - chủ), do vậy nếu máy chủ chứa cơ sở dữ liệu bị tấn công gây mất mát dữ liệu, dữ liệu trên Cổng dữ liệu mở sẽ không còn để sẵn sàng sử dụng. Ngoài ra quan trọng hơn là nếu máy chủ của hệ thống Cổng dữ liệu mở bị tấn công, hệ thống sẽ bị tạm dừng hoạt động. Đây cũng chính là nhược điểm chung của kiến trúc mô hình khách - chủ (Client - Server) hiện nay.

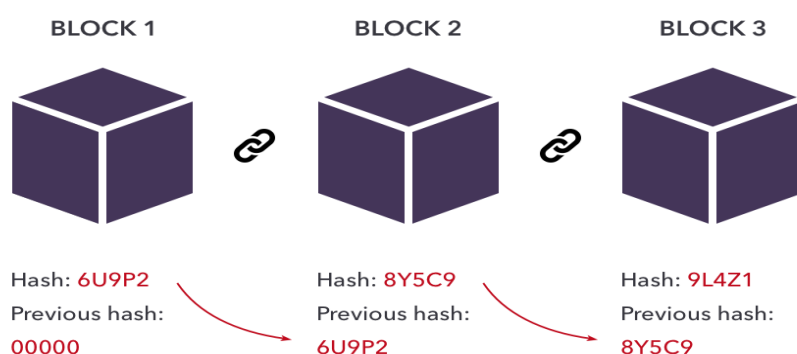
Việc đảm bảo bảo mật cho hệ thống Cổng dữ liệu mở là yêu cầu bắt buộc vì dù dữ liệu được công khai, nhưng những người không có quyền thay đổi dữ liệu như Quản trị hệ thống có thể can thiệp trực tiếp vào cơ sở dữ liệu và thay đổi nội dung trên hệ thống cổng dữ liệu mở. Nếu không có cơ chế bảo mật phù hợp, cổng dữ liệu mở có thể bị thay đổi dữ liệu trên hệ thống.

Do đó, cần xây dựng một kiến trúc hệ thống có khả năng đảm bảo an toàn dữ liệu sao cho dữ liệu không thể bị can thiệp hoặc sửa đổi trái phép, đồng thời dữ liệu có khả năng phân tán trên nhiều node mạng (máy chủ). Khi một node nào đó gặp sự cố thì tất cả các node còn hoạt động còn lại vẫn có thể hoạt động bình thường. Node gặp sự cố sau khi được khắc phục có thể tự khôi phục và cập nhật dữ liệu mới từ hệ thống.

1.3. Giới thiệu công nghệ Blockchain

Blockchain là một cơ sở dữ liệu phân tán (Distributed Ledger), trong đó dữ liệu được lưu trữ thành các khối (Block). Mỗi khối chứa thông tin dữ liệu, mã băm (hash) của chính nó và mã băm của khối trước đó. Các khối được liên kết với nhau thông qua giá trị băm này, tạo thành một chuỗi liên tục gọi là chuỗi khối (Blockchain) [1], [2].

Nếu một khối trong Blockchain bị thay đổi nội dung, giá trị băm của khối đó sẽ thay đổi theo, làm sai lệch liên kết với các khối phía sau. Do đó, toàn bộ chuỗi sau khối bị sửa đổi sẽ không còn hợp lệ.



Hình 1. Cấu trúc chuỗi khối (Blockchain) với cơ chế liên kết băm [1]

Ngoài ra, dữ liệu của Blockchain sẽ được phân tán ở tất cả các node tham gia vào hệ thống Blockchain, khi có các nodes bị lỗi thì hệ thống vẫn có thể hoạt động bình thường miễn là có một số node không bị lỗi. Khi node bị lỗi hoạt động trở lại, hệ thống sẽ tự động khôi phục và cập nhật dữ liệu cho node đó.

2. PHƯƠNG PHÁP/ VẬT LIỆU NGHIÊN CỨU

2.1. Giao thức đồng thuận

Trong hệ thống Blockchain, giao thức đồng thuận là cơ chế cho phép các nút trong mạng đạt được sự thống nhất trong việc tạo khối chứa dữ liệu (block) mới [3]. Cơ chế này đóng vai trò bảo đảm tính phi tập trung hoá của hệ thống Blockchain. Tùy theo mục tiêu thiết kế và môi trường triển khai, các hệ thống Blockchain lựa chọn những giao thức đồng thuận khác nhau. Có hai loại giao thức đồng thuận chính sau:

Proof of Work (PoW - dựa trên bằng chứng công việc) là một trong những cơ chế đồng thuận đầu tiên, được áp dụng trong Bitcoin [4]. Trong mô hình này, các nút tham gia phải thực hiện việc tính toán một bài toán khó phức tạp nhằm tìm ra lời giải hợp lệ để giành quyền tạo block mới (qua đó nhận được tiền thưởng của hệ thống). Ưu điểm của PoW là mức độ bảo mật cao và công bằng, ai đó đầu tư năng lực xử lý càng nhiều thì năng lực tính toán càng mạnh, khả năng giải được bài toán khó trong thời gian lớn càng cao. Tuy nhiên, nhược điểm của cơ chế này là tiêu tốn năng lượng đáng kể, tốc độ xử lý giao dịch thấp vì cần phải giải được bài toán khó mới có thể tạo được block. Ngoài ra điểm đặc biệt quan trọng là để đảm bảo mức an toàn cao, hệ thống phải có đủ số lượng lớn các nodes tham gia. Các nhược điểm này không phù hợp với các hệ thống yêu cầu hiệu suất cao và có ít nodes tham gia trong môi trường tổ chức hoặc hành chính công.

Proof of Stake (PoS - dựa trên bằng chứng cổ phần) là cơ chế đồng thuận ra mắt sau PoW với mục tiêu khắc phục một số nhược điểm của PoW. PoS hoạt động dựa trên nguyên tắc lựa chọn node giành quyền tạo block mới theo lượng tài sản đặt cọc (cổ phần) của node đó. Cụ thể, với mỗi lần tạo block mới, giải thuật PoS sẽ chọn ra người có quyền tạo block dựa trên tính chất rằng nếu ai đó có nhiều cổ phần trong hệ thống thì sẽ có xác suất trở thành người có quyền tạo block cao hơn. So với PoW, PoS giúp giảm đáng kể mức tiêu thụ năng lượng (do không phải đi giải bài toán khó) và cải thiện hiệu suất xử lý giao dịch.

2.2. Đề xuất giao thức đồng thuận cho công dữ liệu mở

Trong hệ thống Công dữ liệu mở tỉnh Thanh Hóa số lượng các node (máy chủ) tham gia và lưu trữ dữ liệu là hạn chế. Do vậy không thể lựa chọn giải thuật đồng thuận PoW do mức an toàn của PoW dựa trên số lượng lớn node tham gia hệ thống. Ngoài ra tương tự như vậy cũng không thể chọn giải thuật đồng thuận PoS dạng public Blockchain (chuỗi khối công cộng) như Algorand, Gasper, do mức an toàn của các giải thuật đồng thuận dạng public Blockchain cũng dựa trên số lượng node tham gia vào hệ thống. Với phân tích ở trên và đặc tính của hệ thống Công dữ liệu mở tỉnh Thanh Hóa, chúng tôi lựa chọn giải thuật đồng thuận PoS dạng private Blockchain (chuỗi khối riêng tư). Trong các giao thức đồng thuận PoS dạng private Blockchain, Raft là giao thức đồng thuận được chúng tôi lựa chọn bởi lý do cốt lõi vấn đề mà thuật toán Raft giải quyết là việc triển khai các máy trạng thái nhân bản (replicated state machines). Với cách tiếp cận này, một tập hợp các máy chủ duy trì các bản sao giống hệt nhau của cùng một trạng thái. Ngay cả khi một số máy chủ gặp sự cố, hệ thống vẫn có thể tiếp tục hoạt động. Để đạt được điều này, một thuật toán đồng thuận sẽ quản lý một nhật ký nhân bản (replicated log). Mỗi máy chủ lưu trữ một bản sao của nhật ký này, chứa một chuỗi các lệnh. Thuật toán đảm bảo rằng mọi nhật ký cuối cùng sẽ chứa cùng một chuỗi lệnh theo cùng một thứ tự. Do đó, mỗi máy trạng thái sẽ xử lý cùng một chuỗi lệnh, tạo ra ảo giác rằng các máy chủ này hoạt động như một thực thể duy nhất, có độ tin cậy cao [5].

2.3. Giao thức đồng thuận Raft

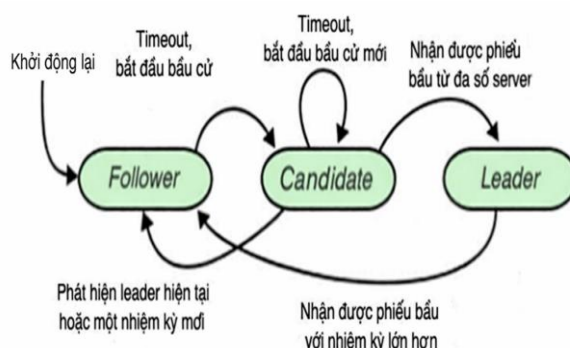
Raft là một thuật toán để quản lý nhật ký được sao chép và thực hiện đồng thuận bằng cách đầu tiên bầu ra một người Leader, sau đó giao cho Leader toàn quyền quản lý nhật ký được sao chép. Leader tiếp nhận các mục nhật ký từ máy khách, sao chép chúng trên các máy chủ khác và thông báo cho máy chủ khi nào an toàn để áp dụng các mục nhật ký vào máy trạng thái của chúng. Leader có thể quyết định vị trí đặt các mục nhật ký mới mà không cần tham khảo ý kiến của các máy chủ khác, và dữ liệu sẽ được truyền tải một cách đơn giản từ người lãnh đạo này sang các máy chủ khác. Nếu Leader gặp sự cố hoặc bị ngắt kết nối với các máy chủ khác thì một người lãnh đạo mới sẽ được bầu [5].

Một cụm Raft chứa nhiều máy chủ, cho phép hệ thống chịu được một số lỗi nhất định. Tại bất kỳ thời điểm nào, mỗi máy chủ đều ở một trong ba trạng thái:

Leader (Người lãnh đạo): là nút duy nhất có quyền xử lý yêu cầu từ máy khách. Mọi yêu cầu gửi đến các nút khác đều được chuyển tiếp về Leader để xử lý tập trung.

Follower (Người theo dõi): là trạng thái thụ động - các nút này không tự khởi xướng bất kỳ hành động nào mà chỉ phản hồi các yêu cầu đến từ Leader hoặc ứng viên. Đây là trạng thái mặc định của phần lớn các nút trong điều kiện hoạt động bình thường.

Candidate (Ứng viên): là trạng thái trung gian, xuất hiện khi một Follower nghi ngờ Leader hiện tại đã ngừng hoạt động và chủ động đề cử bản thân để bầu chọn Leader mới.



Hình 2. Trạng thái của các Node trong Raft [5]

Trong Raft, chu kỳ được chia thành các nhiệm kỳ có độ dài tùy ý và được đánh số bằng các số nguyên liên tiếp. Mỗi nhiệm kỳ được bắt đầu bằng một cuộc bầu cử. Nếu một node nào đó trở thành Leader, Leader đó sẽ giữ vai trò trong suốt thời gian của nhiệm kỳ. Trong một số trường hợp, cuộc bầu cử sẽ không thể lựa chọn được Leader thì nhiệm kỳ sẽ kết thúc mà không có người lãnh đạo; một nhiệm kỳ mới (với một cuộc bầu cử mới) sẽ sớm bắt đầu. Raft đảm bảo chỉ có tối đa một người lãnh đạo trong một nhiệm kỳ.

3. KẾT QUẢ NGHIÊN CỨU VÀ THẢO LUẬN

3.1. Đề xuất kiến trúc hệ thống Cổng dữ liệu mở tích hợp Blockchain

Để giải quyết các thách thức về tính toàn vẹn và bảo mật của cơ sở dữ liệu truyền thống, chúng tôi đề xuất một kiến trúc hệ thống cổng dữ liệu mở lai kết hợp giữa hệ thống CSDL tập trung hiện có và hệ thống Blockchain. Kiến trúc tổng thể của hệ thống được chia thành 3 lớp (layer) hoạt động độc lập nhưng liên kết chặt chẽ với nhau [6]:

Lớp ứng dụng: Đây là giao diện tương tác trực tiếp (Web Portal) dành cho người dân, doanh nghiệp và các cán bộ quản trị. Lớp này tiếp nhận các yêu cầu truy vấn dữ liệu, tải lên bộ dữ liệu mới hoặc kiểm tra tính xác thực của thông tin.

Lớp Lưu trữ ngoài chuỗi (Off-chain Storage Layer): Do đặc thù dữ liệu mở thường có dung lượng lớn, việc lưu trữ toàn bộ nội dung lên Blockchain là không khả thi và gây giảm hiệu suất hệ thống. Do đó, các tệp dữ liệu vật lý (tài liệu, hình ảnh, file CSV/Excel) vẫn sẽ được lưu trữ tại máy chủ cơ sở dữ liệu truyền thống của cơ quan nhà nước và sử dụng các kỹ thuật sao lưu dữ liệu thông thường.

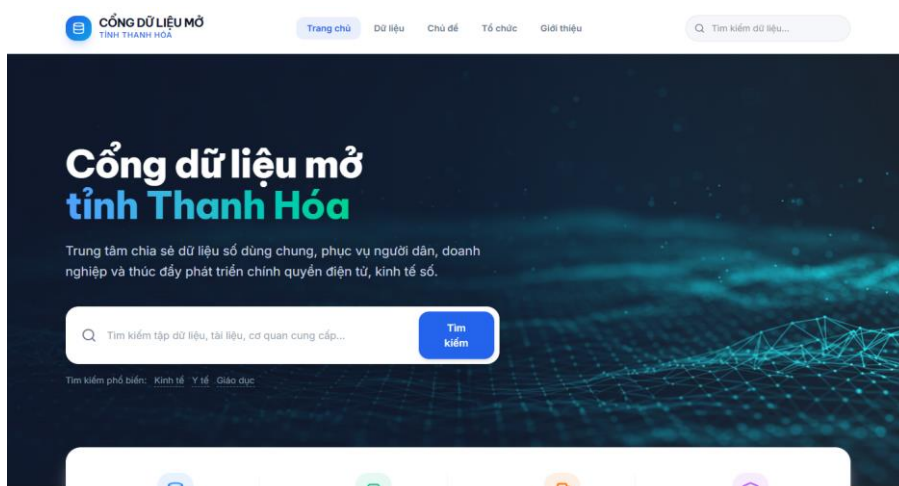
Lớp Blockchain (On-chain/Blockchain Layer): Sử dụng mạng lưới Private Blockchain dựa trên Hyperledger Fabric dùng giao thức đồng thuận Raft. Thay vì lưu dữ liệu gốc, hệ thống chỉ trích xuất và lưu trữ giá trị băm (hash) của dữ liệu cùng các thông tin siêu dữ liệu (metadata) như nguồn gốc, thời gian tạo lên chuỗi khối.

Quá trình đảm bảo an toàn thông tin diễn ra thông qua các hợp đồng thông minh (chaincode) được lập trình trên Hyperledger Fabric [7]. Khi quản trị viên công bố một bộ dữ liệu mở mới, hệ thống lớp ứng dụng sẽ tính toán hàm băm mã hóa của tệp dữ liệu đó. Giá trị băm này ngay lập tức được gửi đến mạng lưới Blockchain. Tại đây, hệ thống sử

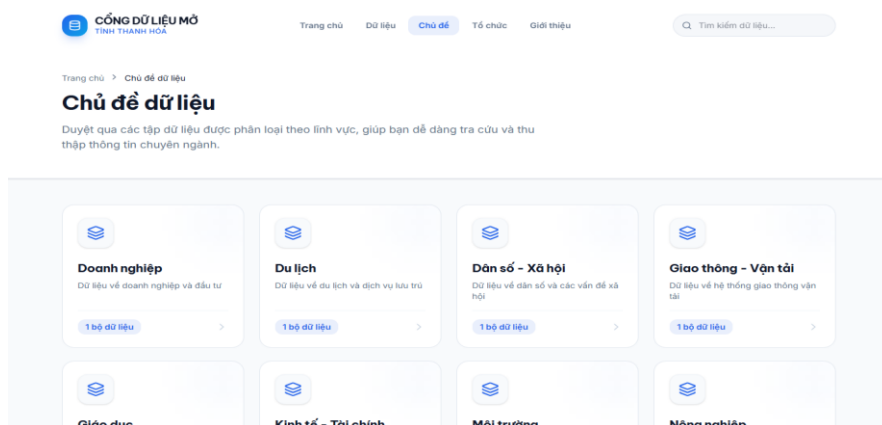
dụng thuật toán đồng thuận Raft để thống nhất trạng thái giữa các máy chủ. Nút lãnh đạo (Leader) sẽ tiếp nhận giá trị băm, sao chép chúng sang các nút theo dõi (Follower) và ghi nhận thành một khối (block) mới không thể đảo ngược. Dữ liệu trên hệ thống sẽ được phân tán ở tất cả các node tham gia, đảm bảo nếu một node gặp sự cố, hệ thống vẫn duy trì hoạt động bình thường. Khi người dùng cuối tải xuống một bộ dữ liệu và muốn xác thực độ tin cậy, hệ thống sẽ tự động băm lại tệp dữ liệu vừa tải và đối chiếu với giá trị băm gốc đang được lưu trên chuỗi khối Blockchain. Bất kỳ sự can thiệp trái phép nào từ tin tặc hoặc quản trị viên hệ thống nhằm thay đổi nội dung tệp dữ liệu đều sẽ làm thay đổi giá trị băm, qua đó hệ thống cảnh báo ngay lập tức về việc dữ liệu đã mất đi tính toàn vẹn. Cơ chế này cho phép truy vết nguồn gốc dữ liệu một cách minh bạch, đáp ứng hoàn toàn yêu cầu khắt khe về an toàn thông tin cho hệ thống cổng dữ liệu mở.

3.2. Chức năng của hệ thống thử nghiệm

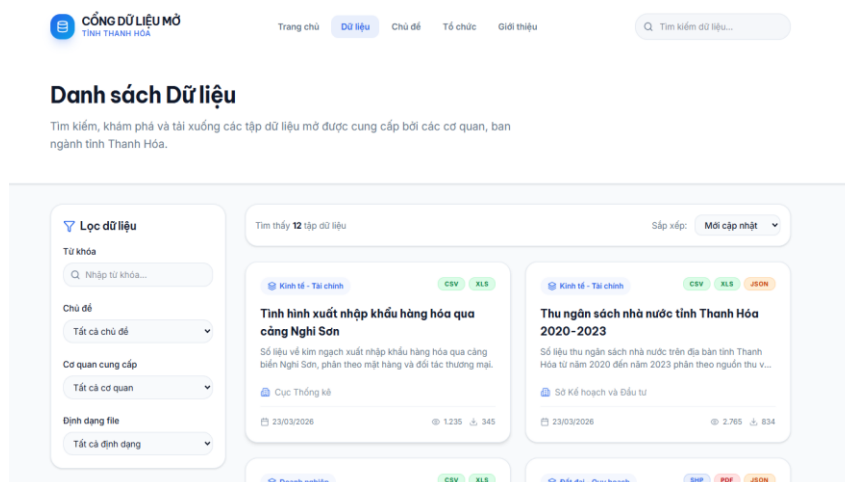
Hệ thống có các chức năng cơ bản như: giúp người dùng tra cứu tập dữ liệu theo từ khoá, theo chủ đề, hiển thị thông tin các tổ chức chủ đề theo đơn vị.



Hình 3. Giao diện cổng dữ liệu mở tỉnh Thanh Hoá thử nghiệm



Hình 4. Giao diện các bộ dữ liệu trong Open Data



Hình 5. Giao diện danh sách các dữ liệu trong cổng dữ liệu mở

3.3. Đánh giá hiệu năng hệ thống

Để đánh giá hiệu năng của hệ thống Cổng dữ liệu mở tích hợp Blockchain, nhóm nghiên cứu đã tiến hành thử nghiệm trên môi trường mô phỏng với cụm Raft gồm 5 node (1 Leader + 4 Follower), sử dụng Hyperledger Fabric phiên bản 2.5. Máy chủ thử nghiệm được cấu hình với CPU Intel Xeon 8 nhân, RAM 32 GB, ổ SSD NVMe 500 GB, kết nối mạng nội bộ 1 Gbps. Các tiêu chí đánh giá bao gồm: thời gian tạo khối (block creation time), độ trễ xác thực dữ liệu (validation latency), và khả năng chịu tải khi có nhiều truy vấn đồng thời (throughput under load).

3.3.1. Thời gian tạo khối (Block Creation Time)

Thời gian tạo khối phản ánh tốc độ ghi nhận dữ liệu lên chuỗi khối. Trong hệ thống Hyperledger Fabric với giao thức Raft, một khối mới được tạo khi đủ điều kiện về số lượng giao dịch hoặc khi hết timeout cấu hình. Thử nghiệm được thực hiện với ba mức tải khác nhau (10, 50, 100 giao dịch/lần) và đo 50 vòng lặp cho mỗi mức. Kết quả được thể hiện trong Bảng 1.

Bảng 1. Thời gian tạo khối theo số lượng giao dịch

Số giao dịch/khối	Thời gian TB (ms)	Thời gian Min (ms)	Thời gian Max (ms)
10	142	118	189
50	318	274	401
100	547	498	623

Kết quả cho thấy với 10 giao dịch/khối, thời gian tạo khối trung bình chỉ khoảng 142 ms - đáp ứng tốt nhu cầu thực tế của hệ thống cổng dữ liệu mở vốn không yêu cầu tần suất ghi dữ liệu cao. Khi tăng lên 100 giao dịch/khối, thời gian tạo khối tăng lên ~547 ms, vẫn nằm trong ngưỡng chấp nhận được so với yêu cầu vận hành. So sánh với giao thức PoW (thời gian tạo khối trung bình từ 10 - 60 giây với Bitcoin), giao thức Raft trong môi trường Private Blockchain thể hiện ưu thế vượt trội về tốc độ tạo khối.

3.3.2. Độ trễ xác thực dữ liệu (Validation Latency)

Độ trễ xác thực là thời gian hệ thống cần để băm lại tệp dữ liệu tải xuống và so sánh với giá trị băm lưu trên chuỗi khối. Thử nghiệm được thực hiện với các tệp dữ liệu có kích thước khác nhau, từ 100 KB đến 50 MB, phản ánh thực tế đa dạng của các bộ dữ liệu mở (CSV, Excel, JSON). Kết quả thể hiện trong Bảng 2.

Bảng 2. Độ trễ xác thực theo kích thước tệp dữ liệu

Kích thước tệp	Thời gian băm SHA-256 (ms)	Thời gian truy vấn Blockchain (ms)	Tổng độ trễ xác thực (ms)
100 KB	3	28	31
1 MB	19	31	50
10 MB	142	33	175
50 MB	689	35	724

Kết quả cho thấy thời gian truy vấn Blockchain gần như không đổi (~28-35 ms) bất kể kích thước tệp, vì hệ thống chỉ lưu và truy vấn giá trị băm cố định 256-bit. Phần biến thiên chủ yếu đến từ thao tác tính toán SHA-256 trên tệp dữ liệu gốc. Với tệp 10 MB trở xuống, tổng độ trễ xác thực dưới 200 ms - hoàn toàn có thể thực hiện theo thời gian thực. Với tệp 50 MB, độ trễ khoảng 724 ms, vẫn ở mức chấp nhận được đối với người dùng cuối trong bối cảnh công dữ liệu mở.

3.3.3. Khả năng chịu tải khi có nhiều truy vấn đồng thời (Throughput)

Thử nghiệm khả năng chịu tải được thực hiện bằng công cụ Apache JMeter, mô phỏng đồng thời nhiều người dùng gửi yêu cầu truy vấn và xác thực dữ liệu lên hệ thống. Kết quả đo throughput (số yêu cầu xử lý thành công mỗi giây - TPS) và tỷ lệ lỗi theo từng mức tải người dùng đồng thời thể hiện trong Bảng 3.

Bảng 3. Khả năng chịu tải của hệ thống theo số lượng người dùng đồng thời

Số người dùng đồng thời	TPS (yêu cầu/giây)	Độ trễ TB (ms)	Độ trễ P95(ms)	Tỷ lệ lỗi (%)
10	87	115	198	0%
50	243	206	387	0%
100	418	239	512	0,3%
200	521	384	891	1,2%
500	534	937	2.140	4,7%

Kết quả cho thấy hệ thống đạt throughput ổn định đến ~534 TPS khi có 500 người dùng đồng thời. Ở mức 100 người dùng, tỷ lệ lỗi chỉ 0,3% với độ trễ trung bình dưới 240 ms - hoàn toàn đáp ứng yêu cầu của một công dữ liệu mở cấp tỉnh. Khi tăng lên 500 người dùng đồng thời, độ trễ P95 tăng lên 2.140 ms và tỷ lệ lỗi đạt 4,7%, cho thấy đây là ngưỡng tải cần xem xét nâng cấp hạ tầng hoặc tối ưu hoá cấu hình Hyperledger Fabric trong tương lai. Nhìn chung, các kết quả thử nghiệm hiệu năng trên lớp Blockchain cho thấy mô hình tích hợp Hyperledger Fabric với giao thức đồng thuận Raft hoàn toàn phù hợp với yêu cầu vận hành thực tế của Công dữ liệu mở tỉnh Thanh Hóa. Hệ thống đạt được sự cân bằng tốt giữa tính an toàn, độ trễ xử lý và khả năng chịu tải, đồng thời vẫn bảo đảm được tính toàn vẹn và khả năng truy vết của dữ liệu công bố.

4. KẾT LUẬN

Ứng dụng công nghệ Blockchain vào kiến trúc công dữ liệu mở là hướng tiếp cận hiệu quả để giải quyết các thách thức về an toàn dữ liệu trong bối cảnh chuyển đổi số. Kết quả phân tích và thử nghiệm cho thấy mô hình đề xuất, sử dụng cơ chế lưu trữ giá trị băm trên chuỗi khối kết hợp giao thức đồng thuận Raft, có khả năng phát hiện kịp thời hành vi sửa đổi dữ liệu trái phép, đảm bảo truy vết nguồn gốc dữ liệu minh bạch và duy trì hoạt động ổn định ngay cả khi một số node gặp sự cố. So với các giao thức đồng thuận khác, Raft vượt trội trong môi trường private Blockchain của cơ quan nhà nước do không đòi hỏi năng lực tính toán lớn và không phụ thuộc cơ chế khuyến khích tiền thưởng. Các kết quả này khẳng định tiềm năng ứng dụng của Blockchain trong việc nâng cao độ tin cậy, tính minh bạch và khả năng chống giả mạo cho Công dữ liệu mở tỉnh Thanh Hóa, góp phần xây dựng nền hành chính số an toàn, hiệu quả.

TÀI LIỆU THAM KHẢO

- [1] Z. Zheng, S. Xie, H. N. Dai, X. Chen, and H. Wang, "Blockchain challenges and opportunities: A survey," *International Journal of Web and Grid Services*, 14(4):352-375, 2018, DOI:10.1504/IJWGS.2018.095647
- [2] M. Swan, *Blockchain: Blueprint for a New Economy*. Sebastopol, CA, USA: O'Reilly Media, 2015.
- [3] D. Yaga, P. Mell, N. Roby, and K. Scarfone, *Blockchain Technology Overview*, National Institute of Standards and Technology, Gaithersburg, MD, USA, NIST IR 8202, 2018.
- [4] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008. [Online]. Available: <https://bitcoin.org/bitcoin.pdf>. Accessed: 21/04/2026.
- [5] D. Ongaro and J. Ousterhout, "In search of an understandable consensus algorithm," in *Proc. 2014 USENIX Annual Technical Conference (USENIX ATC '14)*, Philadelphia, PA, USA, 2014, pp.305-319.
- [6] X. Xu et al., "A taxonomy of blockchain-based systems for architecture design," in *Proc. 2017 IEEE International Conference on Software Architecture (ICSA)*, Gothenburg, Sweden, 2017, pp.243-252.
- [7] K. Christidis and M. Devetsikiotis, "Blockchains and smart contracts for the Internet of Things," *IEEE Access*, vol.4, pp.2292-2303, 2016, DOI:10.1109/ACCESS.2016.2566339
- [8] E. Androulaki et al., "Hyperledger Fabric: A distributed operating system for permissioned blockchains," in *Proc. 13th EuroSys Conference (EuroSys '18)*, Porto, Portugal, 2018, pp. 1-15, DOI:10.1145/3190508.3190538.

Tuyên bố về xung đột lợi ích: Các tác giả tuyên bố không có xung đột lợi ích.

Thông tin tài trợ (nếu có): Nghiên cứu này không nhận được tài trợ từ bên ngoài.