# **THÔNG TIN CHUNG CỦA NHÓM**

* Link YouTube video của báo cáo (tối đa 5 phút):   
  *(ví dụ: https://www.youtube.com/watch?v=AWq7uw-36Ng)*
* Link slides (dạng .pdf đặt trên Github của nhóm):   
  *(ví dụ: https://github.com/mynameuit/CS519.M1.KHCL/TenDeTai.pdf)*
* *Mỗi thành viên của nhóm điền thông tin vào một dòng theo mẫu bên dưới*
* *Sau đó điền vào Đề cương nghiên cứu (tối đa 5 trang), rồi chọn Turn in*

|  |  |
| --- | --- |
| * Họ và Tên: Hà Kiệt Hùng * MSSV: 220101031   1664416206857 | * Lớp: CS2205.APR2023 * Tự đánh giá (điểm tổng kết môn): 9/10 * Số buổi vắng: 1 * Số câu hỏi QT cá nhân: 3 * Số câu hỏi QT của cả nhóm: 3 * Link Github: https://github.com/HaKietHung2905 |

# **ĐỀ CƯƠNG NGHIÊN CỨU**

|  |
| --- |
| **TÊN ĐỀ TÀI (IN HOA)**  NGHIÊN CỨU VÀ KHẢO SÁT HỆ THỐNG CƠ SỞ DỮ LIỆU TIẾT KIỆM NĂNG LƯỢNG. |
| **TÊN ĐỀ TÀI TIẾNG ANH (IN HOA)**  ENERGY - EFFICIENT DATABASE SYSTEMS: A SYSTEMATIC SURVEY |
| **TÓM TẮT** *(Tối đa 400 từ)*  Ngày nay, nhu cầu xây dựng một hệ thống cơ sở dữ liệu hiệu mang lại hiệu quả năng lượng ngày càng tăng, các nghiên cứu đã không ngừng phát triển và đưa ra rất nhiều mô hình khác nhau để đạt được mục tiêu đó. Nghiên cứu này sẽ tập trung vào việc khảo sát cũng như đánh giá các mô hình dữ liệu hiện có và đề xuất chi tiết về một số các kỹ thuật quản lý năng lượng trong cơ sở dữ liệu. Ngoài ra, nghiên cứu còn đưa ra những thách thức cùng những vấn đề còn mở trong lĩnh vực này.  Những đóng góp chính của bài viết này sẽ gồm 3 nội dung:   1. Khảo sát chi tiết về những mô hình đã tồn tại trước đây. 2. Đề xuất và đưa ra mô hình tiết kiệm năng lượng. 3. Nêu ra thách thức và vấn đề trong tương lai. |
| **GIỚI THIỆU** *(Tối đa 1 trang A4)*  Ngày nay, việc tiêu thụ năng lượng đang trở thành một thách thức lớn trong các ngành công nghiệp, đặc biệt là trong lĩnh vực công nghệ thông tin. Tiêu thụ năng lượng không chỉ tạo ra gánh nặng kinh tế đối với doanh nghiệp và chính phủ, mà còn gây tác động tiêu cực đến môi trường tự nhiên. Để giải quyết vấn đề này, nhiều nghiên cứu đã tìm kiếm các giải pháp tiết kiệm năng lượng và xây dựng hệ thống cơ sở dữ liệu hiệu quả. Mặc dù những nghiên cứu này đã phần nào giải quyết được những vấn đề được đặt ra nhưng bên cạnh đó vẫn còn tồn đọng nhiều khuyết điểm chưa được giải quyết triệt để.  Trong quá trình nghiên cứu, chúng tôi nhận ra rằng thay vì giảm khả năng mở rộng của hệ thống như các thiết kế truyền thống trước đây, mô hình mới sẽ tập trung vào việc giảm thiểu chi phí năng lượng trên mô hình hiện có mà không làm ảnh hưởng đến hiệu suất của hệ thống. Điều này mở ra một sự thay đổi quan trọng trong suy nghĩ và phương pháp tiếp cận trong lĩnh vực này.  Trước tiên, chúng tôi thực hiện một khảo sát chi tiết về các mô hình hiện có trong lĩnh vực này. Chúng tôi đánh giá các công trình nghiên cứu đã được công bố trước đó để hiểu rõ những tiến bộ đã đạt được và các phương pháp được áp dụng.  Sau đó, chúng tôi đề xuất một phân loại chi tiết về các kỹ thuật quản lý năng lượng trong cơ sở dữ liệu. Chúng tôi sắp xếp các kỹ thuật này theo nguyên tắc và phương pháp cốt lõi, và mô tả chi tiết về cách chúng hoạt động và tác động đến tiêu thụ năng lượng và hiệu suất.  Cuối cùng, chúng tôi đặt ra một số thách thức và vấn đề còn tồn tại trong lĩnh vực xây dựng hệ thống cơ sở dữ liệu hiệu quả về năng lượng. Bên cạnh đó, chúng tôi cũng đề xuất các hướng tiếp cận và giải pháp tiềm năng để nâng cao hiệu suất và hiệu quả năng lượng trong hệ thống cơ sở dữ liệu.  Thông qua việc trình bày chi tiết vấn đề, chúng tôi hy vọng rằng nghiên cứu này sẽ góp phần vào việc xây dựng những hệ thống cơ sở dữ liệu tiết kiệm năng lượng và tạo ra sự nhận thức về tầm quan trọng của việc quản lý năng lượng trong lĩnh vực công nghệ thông tin. |
| **MỤC TIÊU** *(Viết trong vòng 3 mục tiêu)*  Mục tiêu của nghiên cứu là cung cấp một cái nhìn toàn diện về xây dựng hệ thống tiết kiệm năng lượng, từ việc tìm hiểu và đánh giá các công trình đã có, phân loại các kỹ thuật quản lý năng lượng để đưa ra những thách thức và vấn đề còn chưa được giải quyết trong lĩnh vực này. Chúng tôi mong muốn nghiên cứu này sẽ trở thành một tài liệu giá trị và tham khảo cho các nhà nghiên cứu và người quan tâm trong lĩnh vực hệ thống tiết kiệm năng lượng. Giúp đọc giả hiểu rõ hơn về cách xây dựng các hệ thống có khả năng tối ưu hóa sử dụng năng lượng, giảm thiểu lãng phí và tạo ra môi trường bền vững. |
| **NỘI DUNG VÀ PHƯƠNG PHÁP**  1/ Nội dung.  Bằng cách tìm kiếm những thông tin liên quan tại Science Direct, ACM Digital Library, Web of Science và Google Scholar để tìm ra những mô hình cơ sở dữ liệu tối ưu nhất. Bên cạnh đó kết hợp các kĩ thuật và phương pháp khác để xây dựng một hệ thống đạt hiệu quả cao hơn như: Power profiling, Energy models, Query optimization, Workload management, Caching and indexing techniques, Hardware enhancements.  2/ Phương pháp:   1. Tìm kiếm thông tin liên quan.   Tìm kiếm thông tin liên quan: Sử dụng các dịch vụ truy xuất văn bản trực tuyến như Science Direct, ACM Digital Library, Web of Science và Google Scholar để thu thập thông tin từ các nguồn uy tín và đa dạng. Điều này giúp tìm ra những mô hình cơ sở dữ liệu tối ưu nhất và các kỹ thuật tiết kiệm năng lượng đã được nghiên cứu.  Xây dựng chuỗi tìm kiếm: Sử dụng hai tập từ khóa (T và S) có cùng nghĩa hoặc liên quan để xây dựng chuỗi tìm kiếm. Việc này giúp tối ưu hóa quá trình truy xuất và tìm kiếm thông tin trong các nguồn tài liệu đã đề cập. Chuỗi tìm kiếm được biểu diễn dưới dạng: (([T1, S1] hoặc [T2, S1] hoặc [T3, S1] hoặc [T4, S1]) và ([T1, S2]) ,…). Đồng thời, tiêu chí đánh giá cần được áp dụng để chọn lọc các bài viết nghiên cứu phù hợp, bao gồm các tiêu chí như độ tin cậy, phần trình bày rõ ràng và tính khả thi của ý tưởng và quá trình nghiên cứu. Và các nghiên cứu cần phải có những thông tin thống kê chi tiết và có thể thực hiện lại bởi người đọc.   1. Phương pháp kĩ thuật.   Bên cạnh đó, nghiên cứu cũng áp dụng một số kỹ thuật và phương pháp để xây dựng hệ thống cơ sở dữ liệu tiết kiệm năng lượng và đạt hiệu quả cao hơn. Các kỹ thuật và phương pháp này bao gồm:   * Power profiling: Phân tích và đo lường tiêu thụ năng lượng của hệ thống để hiểu và đánh giá hiệu suất năng lượng. Việc này thường được thực hiện bằng cách sử dụng các thiết bị đo năng lượng hoặc cảm biến năng lượng để đo và ghi lại tiêu thụ năng lượng của hệ thống. Thông tin này giúp xây dựng và lựa chọn các mô hình năng lượng phù hợp cho các hệ thống cụ thể. * Energy models: Xây dựng các mô hình về tiêu thụ năng lượng để dự đoán và đánh giá hiệu quả năng lượng của hệ thống. Trong một máy chủ cơ sở dữ liệu, năng lượng (E) tiêu thụ để chạy một công việc trong một khoảng thời gian (T) có thể được tính toán theo công thức sau: E = P × T, trong đó P là công suất trung bình trong quá trình xử lý công việc, được đo bằng Watt, và T là khoảng thời gian tính bằng giây. Năng lượng (E) được tính bằng đơn vị Joule. * Query optimization: Tối ưu hóa truy vấn cơ sở dữ liệu là quá trình tìm cách cải thiện hiệu suất và giảm tiêu thụ năng lượng của các truy vấn. Có hai nhóm tối ưu hóa chính: tối ưu hóa trên từng truy vấn và tối ưu hóa trên từng khối truy vấn. Trong tối ưu hóa trên từng truy vấn, các mô hình hiện có thường sử dụng các phép đo như số lượng các hoạt động cơ bản (như xử lý bởi CPU và I/O) để tính toán chi phí năng lượng cho mỗi truy vấn. Các mô hình này tập trung vào các thành phần phần cứng quan trọng như CPU, ổ đĩa và bộ nhớ. Một số mô hình cũng xem xét tác động của bộ nhớ cache đến tiêu thụ năng lượng của truy vấn. Các mô hình này cung cấp các công thức tính toán tiêu thụ năng lượng và thường được xây dựng dựa trên các phép đo và thực nghiệm trên hệ thống cụ thể. Trong tối ưu hóa trên từng khối truy vấn, mục tiêu là tối ưu hóa các khối truy vấn, trong đó có nhiều giao dịch chứa các truy vấn được thực thi đồng thời. Điều này đòi hỏi xem xét tác động của các truy vấn cùng nhau và cách tối ưu hóa tại cấp độ khối truy vấn. Việc tối ưu hóa trên từng khối truy vấn có thể giúp giảm tiêu thụ năng lượng và tăng hiệu suất toàn bộ hệ thống. * Workload management: Quản lý tải công việc trong hệ thống để phân phối và tối ưu hóa việc sử dụng năng lượng. Một số ý tưởng và phương pháp có thể áp dựng như: Quản lý năng lượng bộ nhớ, phân phối công việc. * Caching and indexing techniques: Sử dụng các kỹ thuật lưu trữ tạm thời và chỉ mục để cải thiện hiệu suất truy cập dữ liệu và giảm tiêu thụ năng lượng. Caching là một kỹ thuật được sử dụng để lưu trữ tạm thời dữ liệu phổ biến hoặc kết quả truy vấn đã được truy cập gần đây trong bộ nhớ nhanh để cung cấp truy cập nhanh hơn cho các truy vấn tương tự trong tương lai. Việc sử dụng caching giúp giảm thời gian truy cập vào cơ sở dữ liệu và giảm tải lưu lượng truy vấn vào hệ thống. Index là một cấu trúc dữ liệu được sử dụng để tăng tốc độ truy vấn bằng cách tạo ra các khóa tìm kiếm và liên kết chúng với vị trí của dữ liệu tương ứng trong cơ sở dữ liệu. Index giúp giảm thời gian tìm kiếm và truy xuất dữ liệu bằng cách tạo ra một cây tìm kiếm có thứ tự hoặc bảng băm để truy cập vào dữ liệu nhanh hơn. |
| **KẾT QUẢ MONG ĐỢI**    Kết quả của quá trình nghiên cứu đạt được như sau:  Tổng quan về tiến bộ và phương pháp: nghiên cứu phải cung cấp cho người đọc một cái nhìn tổng quan về tiến bộ và các phương pháp đã được sử dụng trong lĩnh vực xây dựng hệ thống cơ sở dữ liệu tiết kiệm năng lượng. Từ việc khảo sát các công trình nghiên cứu đã được công bố, người đọc có thể nhận thức được sự phát triển và các hướng tiếp cận đang được sử dụng trong lĩnh vực này.  Phân loại chi tiết về kỹ thuật quản lý năng lượng: Bài báo đề xuất một phân loại chi tiết về các kỹ thuật quản lý năng lượng trong cơ sở dữ liệu. Phân loại này giúp người đọc hiểu rõ hơn về cách tổ chức và nhóm các kỹ thuật theo các nguyên tắc và phương pháp cốt lõi. Thông qua phân loại này, người đọc có thể nhìn thấy cách các kỹ thuật này hoạt động và tương tác với nhau trong quá trình quản lý năng lượng trong hệ thống cơ sở dữ liệu.  Thách thức và hướng phát triển: Ngoài ra, bài báo cũng chỉ ra những thách thức và vấn đề còn tồn tại trong lĩnh vực xây dựng hệ thống cơ sở dữ liệu tiết kiệm năng lượng. Việc nhấn mạnh các thách thức này giúp người đọc nhận biết các khía cạnh chưa được giải quyết hoặc cần nghiên cứu sâu hơn. Đồng thời, nó tạo động lực và hướng phát triển cho các nghiên cứu và công việc tiếp theo trong lĩnh vực này. |
| **TÀI LIỆU THAM KHẢO** *(Định dạng DBLP)*  [1] Energy Efficiency: The New Holy Grail of Data Management Systems Research. In arXiv: https://arxiv.org/abs/0909.1784  [2]Tom Bostoen, Sape Mullender, and Yolande Berbers. 2013. Power-reduction techniques for data-center storage systems. ACM Computing Surveys  (CSUR) 45, 3 (2013), 1–38  [3] Zichen Xu, Yi-Cheng Tu, and Xiaorui Wang. 2013. Dynamic energy estimation of query plans in database systems. In 2013 IEEE 33rd International  Conference on Distributed Computing Systems. IEEE, 83–92: https://doi.org/10.1109/icdcs.2013.21  [4] Binglei Guo, Jiong Yu, Bin Liao, Dexian Yang, and Liang Lu. 2017. A green framework for DBMS based on energy-aware query optimization and  energy-efficient query processing. Journal of Network and Computer Applications 84 (2017), 118–130: https://doi.org/10.1016/j.jnca.2017.02.015 |

*----- Trang này cố tình để trống -----*