

① Big Data

- big data (cả tạp bộ dữ liệu có kích thước lớn và phức tạp) nhanh và cần thời gian để dạy dỗ nó để có thể quan lý được - không thể bị bắt và xử lý hiện tại
- Dữ liệu (3V và 4V, 5V)
- Volume (kho lượn)
- Velocity (tốc độ)
- Variety (Đa dạng)
- Veracity (đúng xác)
- Value (giá trị)

VD: dữ liệu từ Facebook, camera giám sát

② Virtualization

- Công nghệ cho phép tạo ra phiên bản ảo của tài nguyên phần cứng như máy chủ, bộ nhớ RAM, bộ xử lý

- Dữ liệu

- Tách biệt (tách rời) và lập logik
- Cho phép nhân bản (tạo ra nhiều bản sao)
- mảng ổ đĩa máy tính (vật lý)

VD: VMWare, VirtualBox, Hyper-V





Search

L



LEGION

③ Cloud

Mô hình cung cấp tài nguyên điện toán (nguồn) để tăng
để sử dụng qua Internet

- Đặc điểm:

- + theo yêu cầu
- + trả tiền theo mức sử dụng
- + truy cập nhanh
- Mô hình để sử dụng
- +aaS : bao tay
- + PaaS : nền tảng
- + SaaS : phần mềm

④ Data warehouse

Ta có thể thấy rằng trong dữ liệu từ nhiều nguồn khác nhau
để phân tích và hỗ trợ quyết định

- Đặc điểm

- + Tích hợp
- + Lập trình
- + Bảo mật

VD: Amazon, Redshift, Snowflake

Học thầy không tay học bạn.



www.huongtrang.net

horizontal slot sharing



vertical slot ngang 1 vý

⑤ Hadoop

Français: mảng ngẫu nhiên do Apache phát triển, dùng để lưu trữ dữ liệu (theo cách rõ ràng) phân tán

Thứ nhất: chia

+ HDFS: lưu trữ dữ liệu

+ Map reduce: xử lý dữ liệu song song

+ Xử lý ảnh

VD: xử lý log web, phân tích dữ liệu thời gian

⑥ HDFS (Hadoop Distributed File System)

Hỗ trợ tên lửa đặc quyền cao, có thể dù lỗi vẫn có block, và nhận bàn ăn toàn

Đặc điểm

+ block mặc định 128 MB / 256 MB

+ Dữ liệu được replicate trên nhiều node (3)

+ Khi nào nó chạy theo chế độ ngẫu (nhé node)

VD: Twitter, big data



LEGRION

⑦ Map reduce

- mô hình phân tán gần như bao gồm:
- Map: định rõ dữ liệu và xử lý sao cho
 - Reduce: tổng hợp và thu gọn

Đặc điểm

- + Tối ưu cho batch processing
 - + từ đây phân rã và xử lý lõi
- VD: Đầu vào là chuỗi liên tục và đầu ra là

⑧ Hadoop Yarn

Ông trùm quản lý và cấp tài nguyên Hadoop

→ Đặc điểm

→ Quản lý CPU, RAM trên toàn cluster

→ Hỗ trợ tối đa nhiều ứng dụng

VD: chạy Spark trên Hadoop thông qua YARN

⑨ HBase

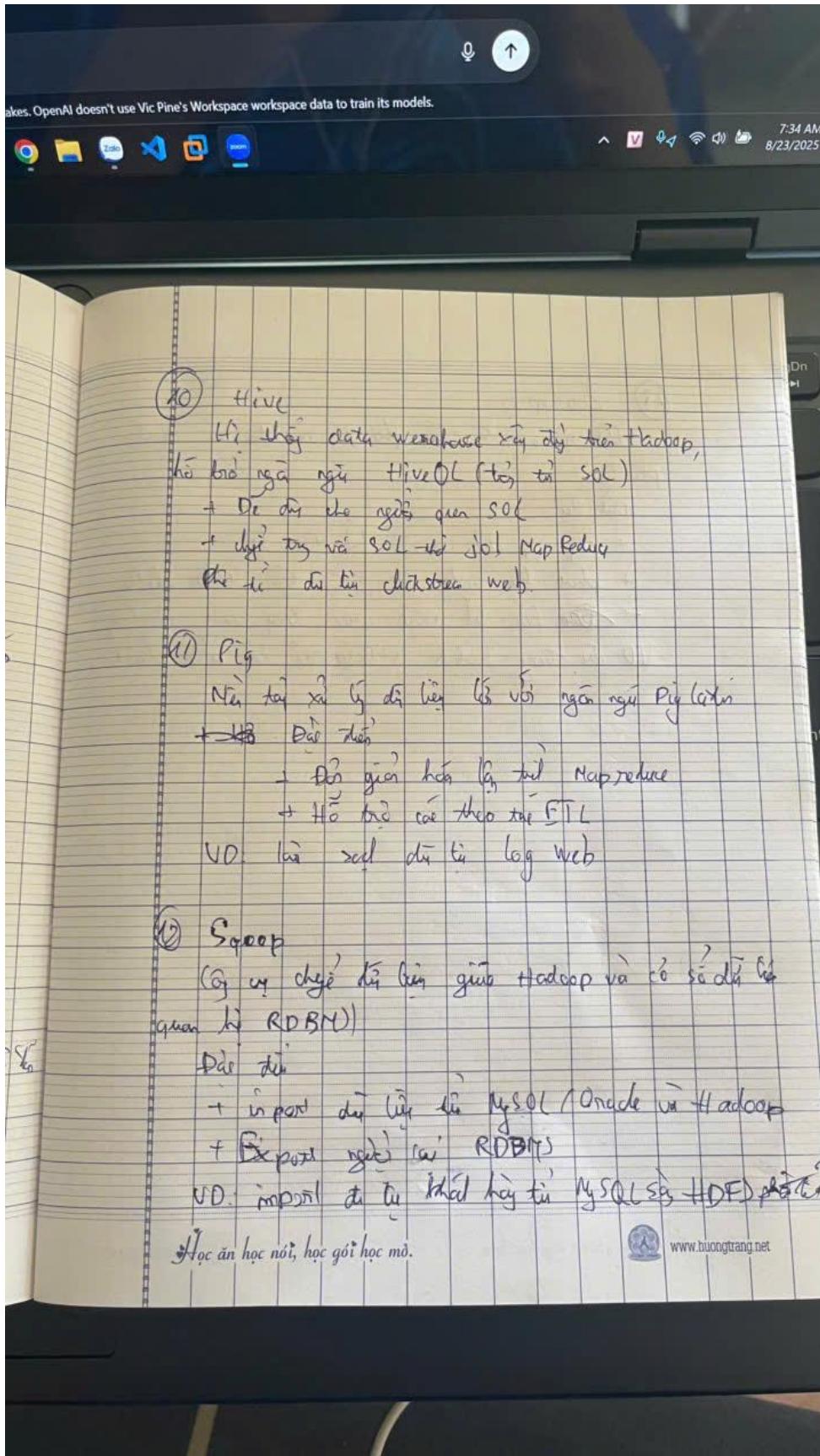
Có thể dùng MySQL, chạy trên HDFS, để lưu trữ

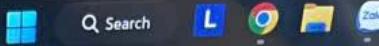
- Nó là:
- ~~SQL~~ NoSQL, key-value, schema-less
 - Tuy vẫn thường gán thuộc tính
 - Tiếp cận dữ liệu phi cấu trúc

Học thầy không tày học bạn. FB: Huynh message
HBase



www.huongtrang.net





⑬ Zookeeper

Dù v. anh kỵ, đều phô bày bùi cho cá áy dày

phô ta

bùi dùi

+ cát cát dìi up cát hít hít chay

+ phôe hít rode lòi, hó - trú keleb delia

+ Bán hít hít khát uyên trọng cua Hadza

VD dù trọng Rath, + Bryce áo đeo phô chita

Học thầy không tày học bạn.



www.hungtran.net

