

www.flex.edu.vn giangdayit@gmail.com

Big Data Fundamentals

MSc. Huynh Nam

BY MR. HUYNH NAM

1

www.flex.edu.vn giangdayit@gmail.com

Nội dung

- Distributed System
- Big Data Fundamentals
- Docker
- Open-Stack
- Hadoop Eco System
- Big Data Analytics

BY MR. HUYNH NAM

2

www.flex.edu.vn giangdayit@gmail.com

Quy định nhóm

- Nhóm chia theo ca thực hành
- Quy mô nhóm theo lớp học
 - Lớp CLC: Mỗi nhóm từ **6 đến 8** sinh viên
 - Lớp đại trà: Mỗi nhóm từ **8 đến 10** sinh viên

BY MR. HUYNH NAM

3

www.flex.edu.vn giangdayit@gmail.com

Thang điểm

- Thang điểm Lý thuyết
 - TK01 = trắc nghiệm
 - TK02 = điểm danh **TK02 < 4.0 sẽ cấm thi cuối kì**
 - MTE (điểm giữa kì) = PRJ (70% báo cáo nhóm) + INV (30% thi đánh giá cá nhân)
 - FE (điểm cuối kì) = PRJ (70% báo cáo nhóm) + INV (30% thi đánh giá cá nhân)
- Thang điểm Thực hành
 - TH01 = PRJ của MTE
 - TH02 = PRJ của FE
 - TH03 = điểm danh **TH03 < 4.0 sẽ cấm thi cuối kì**
- **Một trong 2 cột PRJ hoặc INV = 0 thì MTE, FE = 0**

BY MR. HUYNH NAM

4

www.flex.edu.vn giangdayit@gmail.com

Hình thức báo cáo

- Biên bản đánh giá làm theo mẫu
 - Tỷ lệ đánh giá toàn dự án 100%
 - Tỷ lệ đóng góp từng thành viên phải được định lượng cụ thể
 - Minh chứng đánh giá tỷ lệ đóng góp thể hiện qua phân công trên Jira
- Tài liệu báo cáo giữa kì:
 - Thời gian: báo cáo theo nhóm thực hành diễn ra vào **buổi 05** theo lịch thực hành
 - Slide (.pptx)
 - Sản phẩm hoàn thiện
- Tài liệu báo cáo cuối kì: Nộp trước ngày báo cáo 1 ngày thông qua email: giangdayit@gmail.com
 - Thời gian: báo cáo theo nhóm thực hành diễn ra vào **buổi 10** theo lịch thực hành
 - Slide (.pptx)
 - Sản phẩm hoàn thiện
 - Report (.docx)
- Toàn bộ source code dự án phải được thể hiện quá trình làm trên GitHub / Gitlab
- Tất cả các thành viên đều tham gia báo cáo (từng cá nhân)
 - Thuyết trình
 - Thực hành code
 - Trả lời câu hỏi
- Vắng mặt báo cáo thì cột điểm **PRJ = 0**

BY MR. HUYNH NAM

5

www.flex.edu.vn giangdayit@gmail.com

Nội dung đề án

- Dự án giữa kì: Xây dựng hệ phân tán
 - Sinh viên tìm hiểu xây dựng docker swarm và quản trị swarm. Triển khai hệ thống phân mềm phân tán trên cụm docker được hiện thực hóa bằng Virtual Box với các máy ảo Linux theo version được quy định.
- Dự án cuối kì: Xây dựng hệ thống cloud cung cấp dịch vụ lưu trữ, xử lý và phân tích dữ liệu lớn
 - Tìm hiểu OpenStack xây dựng private cloud trên môi trường ảo hóa Virtual Box với cấu hình, version được quy định.
 - Xây dựng cụm Kubernetes và Pod trên private cloud.
 - Sử dụng docker container trên cụm Kubernetes để triển khai và vận hành hệ thống Hadoop phục vụ Big Data Analytics với MapReduce, Pig, Hive và Spark.

BY MR. HUYNH NAM

6

www.flex.edu.vn giangdayit@gmail.com



THANK YOU
Q & A

BY MR. HUYNH NAM

7