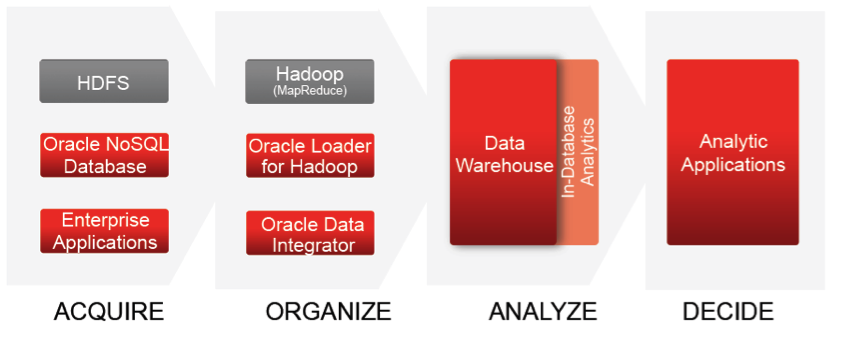
**BIG DATA**

# Giới thiệu

Để có thể nghiên cứu về Big Data, trước tiên cần phải có một nguồn dữ liệu thật lớn. Sau đó phải có chỗ để lưu trữ nó. Trên thế giới hiện nay có rất nhiều giải pháp cho mảng Big Data, mỗi giải pháp giải quyết một bài toán, yêu cầu khác nhau. Nhưng tựu chung lại có thể được chia thành các thành phần chính trong bức tranh tổng thể về big data như sau:



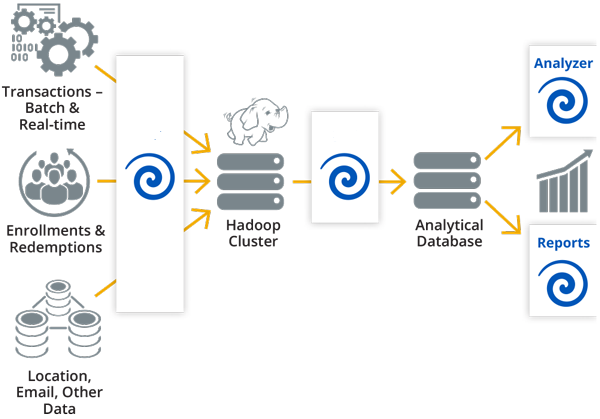
* Thu thập dữ liệu (ACQUIRE): Ở bước này, **ELCOM phải lựa chọn công nghệ và giải pháp để thu thập và lưu trữ dữ liệu**. Với các hãng công nghệ lớn thì có các giải pháp như: Oracle NoSQL Database, Amazon AWS có giải pháp Amazon S3, Google thì có Google BigTable. Dòng sản phẩm nguồn mở thì có các công nghệ tiêu biểu như: MongoDB, Cassandra, Apache Hbase.
* Tổ chức dữ liệu (ORGANIZE): Ở bước tổ chức, tức là có thể xử lí dữ liệu lớn phân tán, song song. Ơr bước này hầu hết các giải pháp hiện nay đều dùng mô hình Map-Reduce. **Do vậy công việc chính của ELCOM ở bước này là lựa chọn ngôn ngữ để implement phần nghiệp vụ xử lý theo mô hình đề xuất của Hadoop.**
* Phân tích dữ liệu (ANALYZE): Bao gồm các giải pháp để phục vụ bài toán Data Warehause**. ELCOM phải tìm hiểu và phát triển giải pháp để đáp ứng các yêu cầu của một tiến trình sao chép và chuyển đổi dữ liệu (ETL).**
* Ra quyết định (DECIDE): Khai phá dữ liệu, trích xuất các thông tin hữu ích (big data -> smart data) để ra các quyết định. Ở giai đoạn này, **ELCOM phải tìm hiểu các thuật toán học máy, xử lý ngôn ngữ tự nhiên, data mining**.

# Mục tiêu 06 tháng cuối năm 2014

Để nghiên cứu và xây dựng hệ thống Big Data cần nghiên cứu và tìm hiểu nhiều công nghệ đặc thù như cơ sở dữ liệu, thuật toán, công nghệ xử lý thời gian thực …

Mục tiêu 06 tháng cuối năm là xây dựng được một khung về mặt công nghệ để có thể tổ **chức lưu trữ, truy cập và phân tích** Big Data và áp dụng để giải quyết cho bài toàn CDR Mediation trong dự án C1. Đối chiếu với mô hình ở trên, ELCOM sẽ đạt được ở mức 3 – Phân tích dữ liệu. Như vậy đối với mảng Big Data tại ELCOM, kết thúc năm 2014 chúng ta có được:

* Cơ sở dữ liệu để phục vụ cho yêu cầu thu thập và lưu trữ dữ liệu lớn.
* Xây dựng được giải pháp implement theo mô hình Map/Reduce và xử lý phân tán.
* Xây dựng được Data Warehause và bộ công cụ ETL.
* Áp dụng vào bài toán CDR Mediation.



**CÔNG NGHỆ XỬ LÝ THỜI GIAN THỰC**

# Giới thiệu

Xử lý dữ liệu theo thời gian thực đã và đang là một yêu cầu cần phải đạt được đối với các ứng dụng chạy trên nền Web, các ứng dụng xử lý dữ liệu và các ứng dụng phân tích nghiệp vụ khác. Đặc biệt hơn nữa, với xu thế của big data hiện nay, yêu cầu về tốc độ xử lý trên các nguồn dữ liệu lớn, có cấu trúc phức tạp ngày càng trở nên cấp thiết.

Hiện nay, trong các dự án của ELCOM, yêu cầu về xử lý thời gian thực đã và đang trở thành một trong những yêu cầu mấu chốt cần phải đạt được. Chính vì vậy yêu cầu đặt ra là phải xây dựng được giải pháp xử lý theo thời gian thực chuẩn cho từng dạng bài toán, dự án cụ thể. Hiện tại có hai mô hình xử lý theo thời gian thực được sử dụng phổ biến đó là:

* + **Non-blocking I/O**: Kỹ thuật này được áp dụng cho các ứng dụng trên nền Web để tối ưu các thao tác vào ra như đọc/ghi, truy cập cơ sở dữ liệu.
  + **In-Memory Processing**: Công nghệ xử lý bằng bộ nhớ trong

# Mục tiêu

**XỬ LÍ ẢNH**

# Tìm hiểu các chức năng, đặc điểm của phần mềm xử lý ảnh hiện tại của Elcom.

* 1. **Các công nghệ:**

Hiện trạng: Vẫn đang tiếp tục nghiên cứu nội bộ chưa đưa vào sản phẩm

1. Nâng cao chất lượng ảnh

* Denoise Image (khử nhiễu)
* Video Stabilization (chống rung)

1. Nhận dạng

* Biển số xe
* Trạm điện: Nhận dạng bật – tắt đèn

1. Tracking

* Trên đường cao tốc (các loại xe)
* Trên biền ( tàu thuyền)

1. Mix ảnh

* Sử dụng trong CCTV, hội nghị truyền hình

1. Hiển thị video (vẽ ảnh)

* Hiện đang sử dụng Direct X – là tập các thư viện của HĐH, hỗ trợ các phần mềm khác cách thức giao tiếp với card màn hình, sound, …
  1. **Sản phẩm sẽ ứng dụng cụ thể:**

1. **Eokit ship** – Camera quan sát trên biển

* Denoise Image: khử nhiễu + khôi phục lại chi tiết
* Video Stabilization: chống rung
* Object tracking: theo dõi đối tượng
  + Đang ở mức độ bước đầu nghiên cứu

1. **Electricity Station –** Hệ thống nhận dạng đèn báo trong trạm điện

* Nhận dạng màu các đèn báo tín hiệu (bật hoặc tắt)
* Đã hoàn thiện, độ chính xác cao do yếu tố môi trường đã cố định.

1. **ITS** – Hệ thống giao thông thông minh

* Đếm số lượng phương tiện: Phương pháp background subtraction kết hợp thuật toán tracking để theo dõi, trích chọn.
* Nhận dạng biển số xe.
* Đo tốc độ các phương tiện: calibrate camera, ước lượng tỉ lệ khoảng cách thực tế của đối tượng trên thực tế so với khoảng cách trên ảnh, từ đó suy ra vận tốc của đối tượng. Hiện tại độ chính xác chưa cao

**Tiêu chí đánh giá hiện tại:**

* Độ chính xác: 85%
* Thời gian xử lí: 0.2- 0.3s ảnh 640x480
* Khối lượng lưu trữ dữ liệu: chưa can thiệp

# Tìm hiểu các đặc điểm, tính năng mới về công nghệ xử lý ảnh của các hãng. Đề xuất các ý tưởng cải tiến phần mềm xử lý ảnh hiện tại của Elcom.

* 1. **Nâng cao chất lượng ảnh:**

Với từng yêu cầu riêng biệt có tiêu chí để đánh giá khác nhau.

* 1. **Các tính năng về công nghệ xử lí ảnh trong giao thông thông minh:**
* Automated Tolling: Thu phí tự động
* Safety Monitoring: Giám sát an toàn
* Road Inspection: Kiếm tra chất lượng đường bộ
* Speed Enforcement: Vi phạm tốc độ
* Red Light Enforcement: Vi phạm đèn đỏ
* Reserved Lane Enforcement: Vi phạm làn đường
* Parking Access: Kiểm soát bãi đỗ xe
* Tunnel Safety: Đường hầm an toàn

**Đề xuất cải tiến sản phẩm hiện có:**

* Thêm tính năng kiểm tra chất lượng đường bộ, vi phạm làn đường.
* Tăng độ chính xác: >90%
* Giảm thời gian xử lí: <100ms ảnh 640x480
* Giảm khối lượng lưu trữ dữ liệu: can thiệp để tinh giảm bộ nhớ

# Lộ trình 6 tháng cuối năm 2014

* 1. **Nâng cao chất lượng ảnh**
* Xây dựng version đầu tiên v1.0
* Hoàn thành các chức năng và tích hợp vào hệ thống Esight
  1. **Hệ thống giao thông thông minh**
* Cải tiến các chức năng đã có như đề xuất
* Mở rộng thêm chức năng
* Xây dựng v1.0 thay thế dần phần mềm đang sử dụng trong ITS