



**UCM**

UNIVERSIDAD CATOLICA DEL MAULE

# *Computación de Alto Rendimiento y Biometría*

<http://www.litrp.cl>

**Laboratorio de Investigaciones Tecnológicas en  
Reconocimiento de Patrones (LITRP)**



**LITRP**

Laboratorio de Investigaciones  
Tecnológicas en Reconocimiento de Patrones  
Universidad Católica del Maule



# Equipo de Trabajo



**LITRP**

Laboratorio de Investigaciones  
Tecnológicas en Reconocimiento de Patrones  
Universidad Católica del Maule

## Investigadores

### Director

Dr. Marco Mora



Dr. Ricardo Barrientos



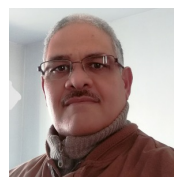
Dr. Ruber Hernández



Dr. David Zabala



Dr. José Naranjo



Dra. Xaviera López



### Áreas de Desarrollo:

- Inteligencia Artificial
- Computación de Alto Rendimiento
- Redes Neuronales
- Tratamiento de Imágenes
- Minería de Datos

### Estudiantes de Postgrado

- Doctorado en Ingeniería
- Doctorado Modelamiento Matemático
- Magíster en Ciencias de la Computación

### Estudiantes de Pregrado

- Ingeniería Civil Informática

# Proyectos financiados por el Gobierno de Chile y la Industria

FONDEF (2012-2013)



FONDEF (2018-2019)



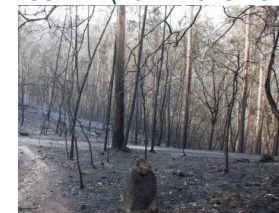
EMPRESA (2017)



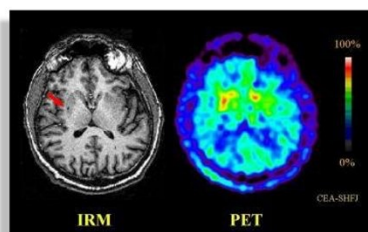
FONDEF (2016-2017)



CONAF (2017-2018-2019)

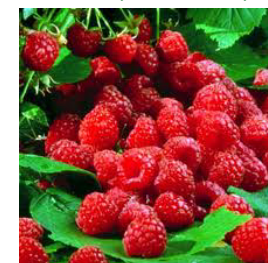


UCM (2017-2018)



- **Propiedad Intelectual (Patentes)**
- **Contratos Tecnológicos con Empresas**
- **Proyectos de Investigación**
- **Publicaciones Internacionales**

FIC-R (2018-2020)





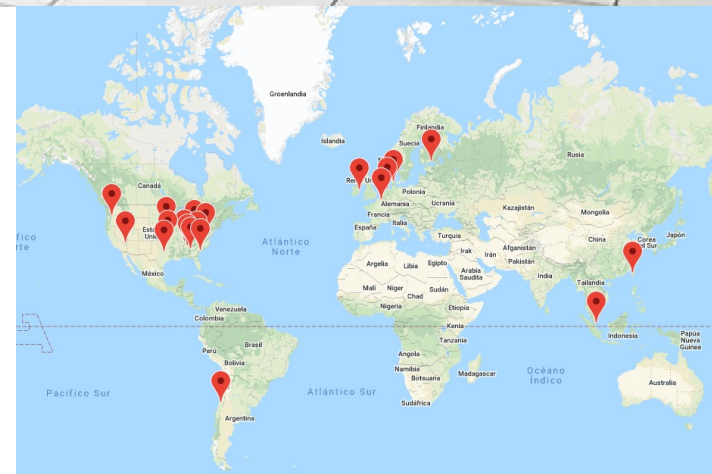
## Ejemplo de problema de Big Data



# Centro de Datos de Google



- Réplicas a lo largo del mundo

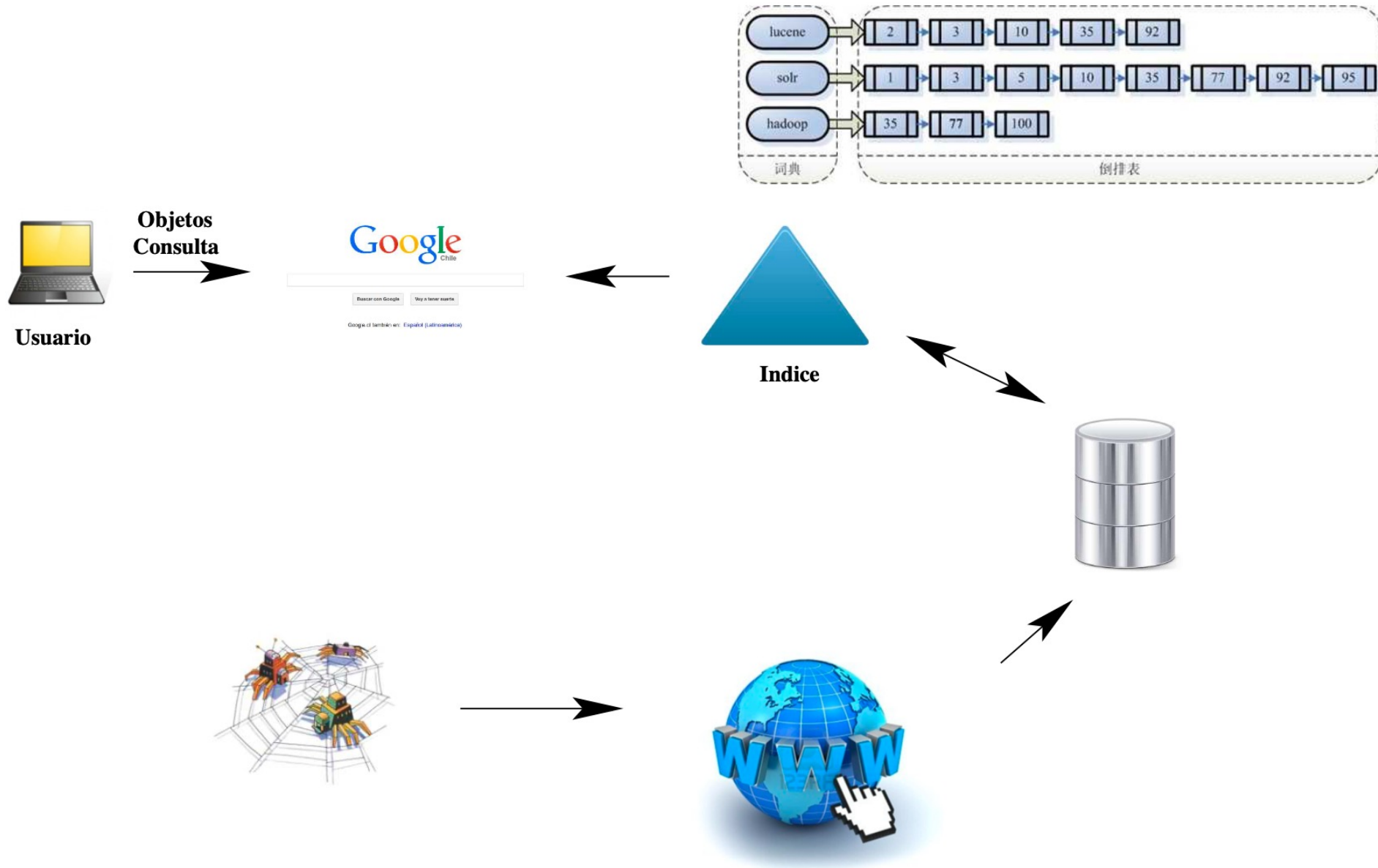


# Planta Fotovoltaica El Romero

- 776.000 módulos solares.
- 1,5 millones de m<sup>2</sup>, equivalente a 211 campos de fútbol.



# Motor de Búsqueda Web





# Computación de Alto Rendimiento

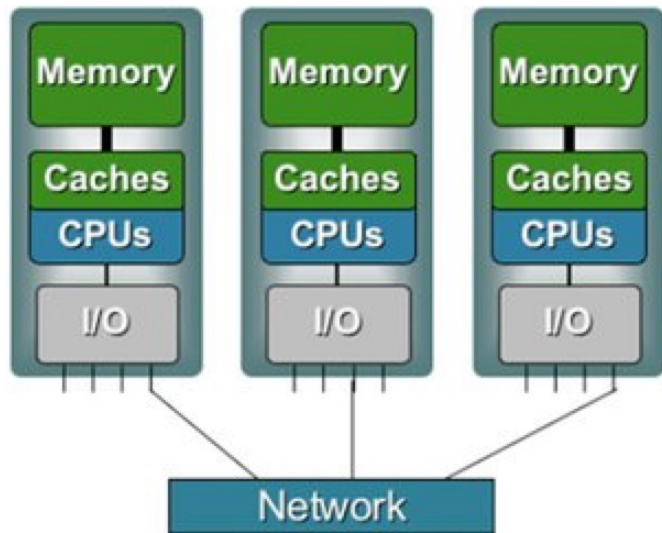
(High Performance Computing, HPC)

[www.top500.org](http://www.top500.org)

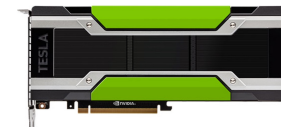
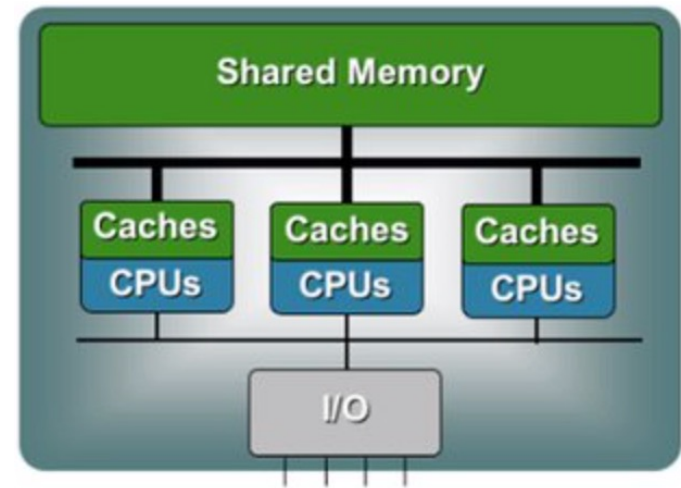
Rank	System	Cores	Rmax (TFlop/s)	Rpeak (TFlop/s)	Power (kW)
1	<b>Supercomputer Fugaku</b> - Supercomputer Fugaku, A64FX 48C 2.2GHz, Tofu interconnect D, Fujitsu RIKEN Center for Computational Science Japan	7,299,072	415,530.0	513,854.7	28,335
2	<b>Summit</b> - IBM Power System AC922, IBM POWER9 22C 3.07GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM DOE/SC/Oak Ridge National Laboratory United States	2,414,592	148,600.0	200,794.9	10,096
3	<b>Sierra</b> - IBM Power System AC922, IBM POWER9 22C 3.1GHz, NVIDIA Volta GV100, Dual-rail Mellanox EDR Infiniband, IBM / NVIDIA / Mellanox DOE/NNSA/LLNL United States	1,572,480	94,640.0	125,712.0	7,438
4	<b>Sunway TaihuLight</b> - Sunway MPP, Sunway SW26010 260C 1.45GHz, Sunway, NRCPC National Supercomputing Center in Wuxi China	10,649,600	93,014.6	125,435.9	15,371
5	<b>Tianhe-2A</b> - TH-IVB-FEP Cluster, Intel Xeon E5-2692v2 12C 2.2GHz, TH Express-2, Matrix-2000, NUDT National Super Computer Center in Guangzhou China	4,981,760	61,444.5	100,678.7	18,482
6	<b>HPC5</b> - PowerEdge C4140, Xeon Gold 6252 24C 2.1GHz, NVIDIA Tesla V100, Mellanox HDR Infiniband, Dell EMC Eni S.p.A. Italy	669,760	35,450.0	51,720.8	2,252
7	<b>Selene</b> - DGX A100 SuperPOD, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100, Mellanox HDR Infiniband, Nvidia NVIDIA Corporation United States	272,800	27,580.0	34,568.6	1,344

# Computación de Alto Rendimiento

*(Sistemas de memoria Distribuida)*



*(Sistemas de memoria compartida)*



# GPU (Graphic Process Unit)

## Año 2008

**100.000.000**  
GPUs aceptan CUDA  
(6.000 son Teslas)

**150.000**  
descargas de CUDA

**1**  
supercomputador  
en el top500.org  
(77 TFLOPS)

**60**  
cursos universitarios

**4.000**  
artículos científicos

## Año 2015

**600.000.000** GPUs aceptan CUDA  
(y 450.000 son Teslas)

**3.000.000** descargas anuales de CUDA  
(una cada 9 segundos)

**75** supercomputadores  
en el TOP500.org  
(acumulado: 54.000 TFLOPS)

**840** cursos universitarios

**60.000** artículos científicos

# GPU (Graphic Process Unit)

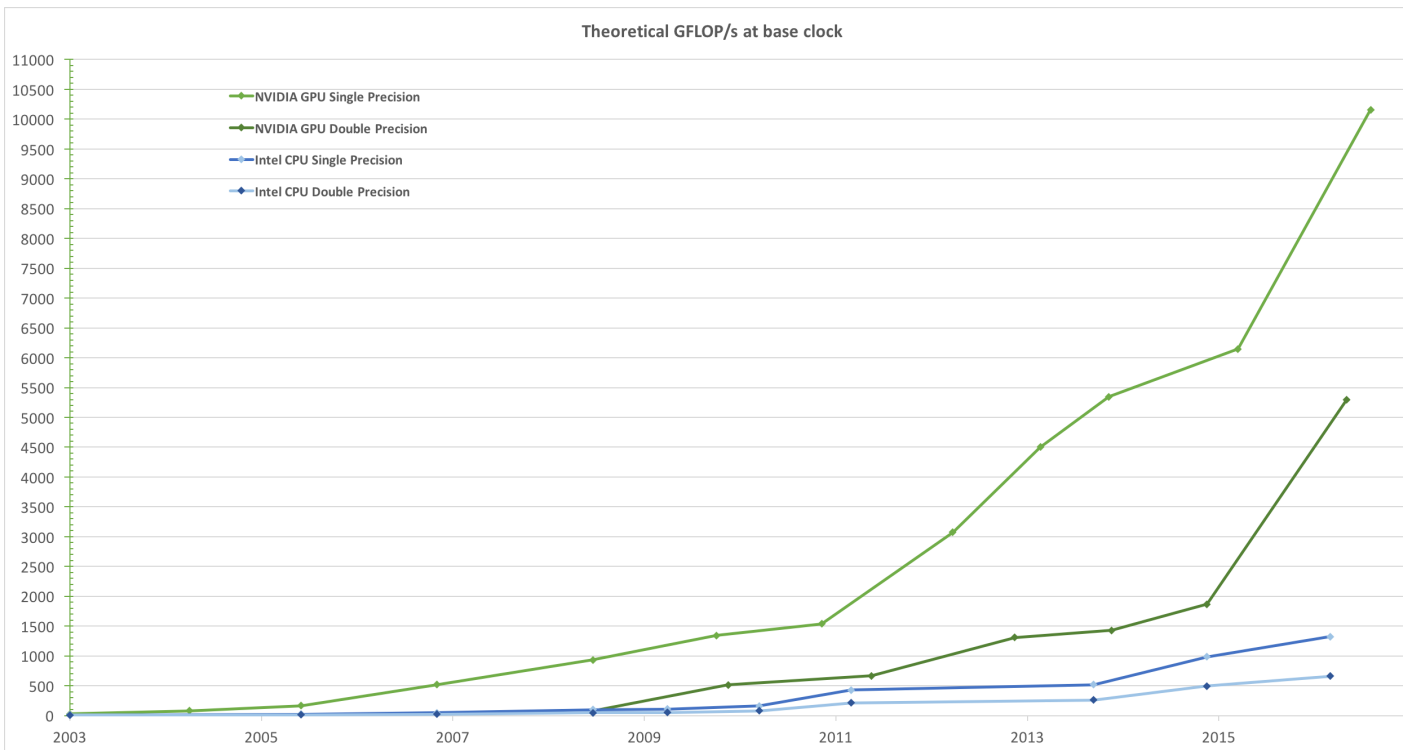


Fig.: CPU vs GPU

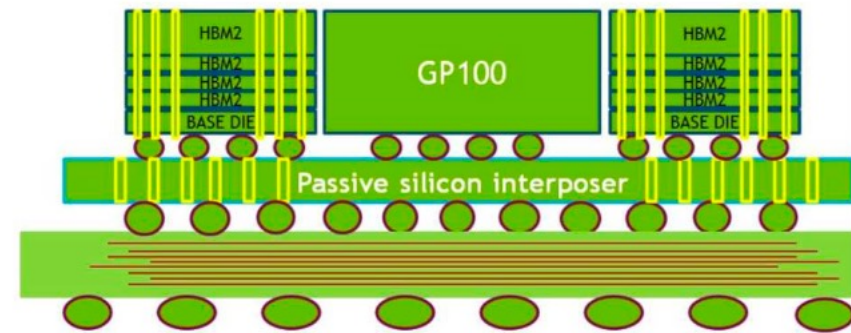
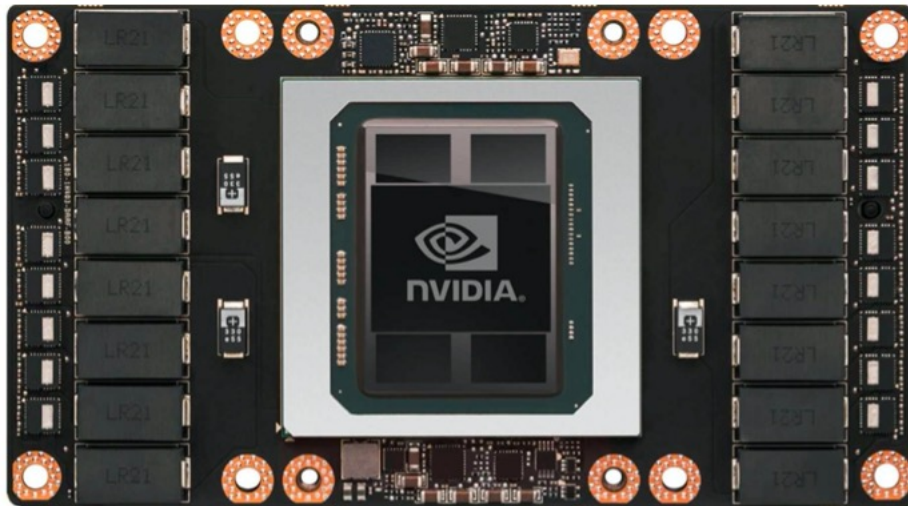


	GeForce 1080Ti
<b>Núcleos</b>	3.584
<b>Memoria</b>	11GB
<b>Precisión Simple</b>	11.340 GFLOPS
<b>Precisión Doble</b>	354 GFLOPS
<b>Banda ancha de memoria</b>	484.4 GB/s

Fig.8: La GPU asigna más transistores al procesamiento de datos.

# GPU (Graphic Process Unit)

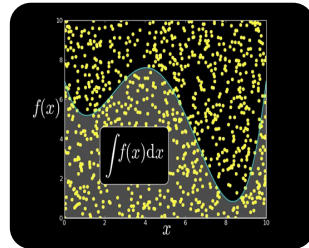
## Memoria HBM2 en GPU Pascal GP100



# Librería aceleradas mediante GPU



NVIDIA cuBLAS



NVIDIA cuRAND



NVIDIA cuSPARSE



NVIDIA NPP



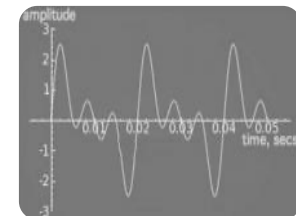
Vector Signal  
Image Processing



GPU Accelerated  
Linear Algebra



Matrix Algebra  
on GPU and  
Multicore



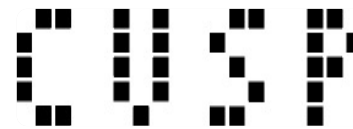
NVIDIA cuFFT



ROGUE WAVE  
SOFTWARE  
IMSL Library



ArrayFire Matrix  
Computations



Sparse Linear  
Algebra

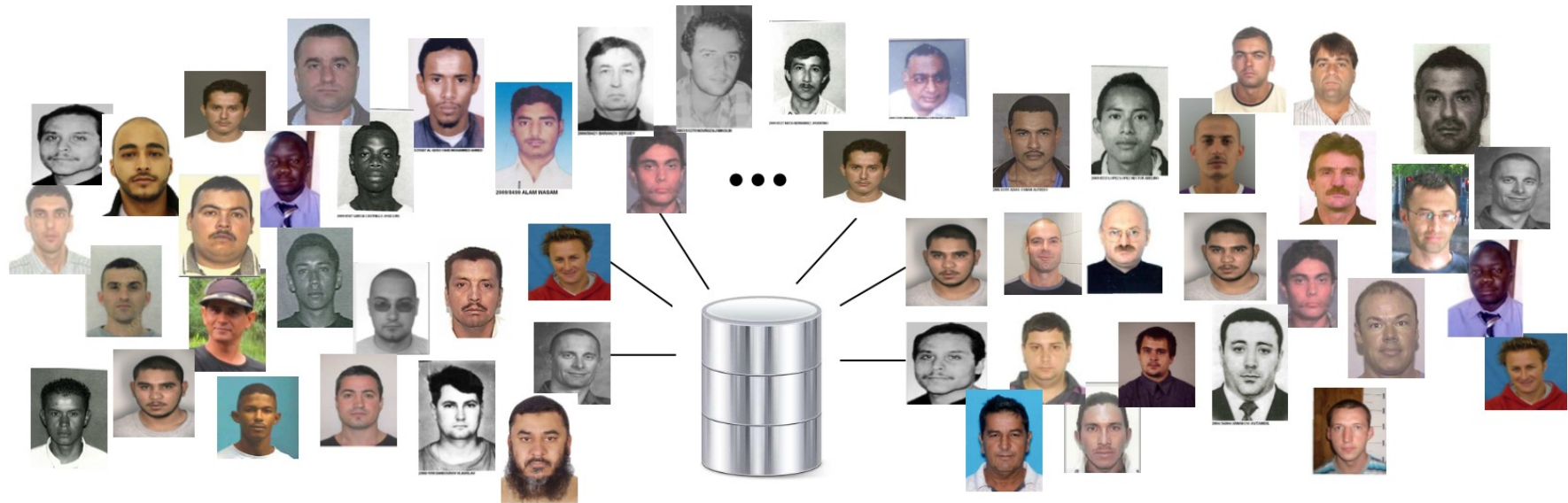


C++ STL  
Features for  
CUDA



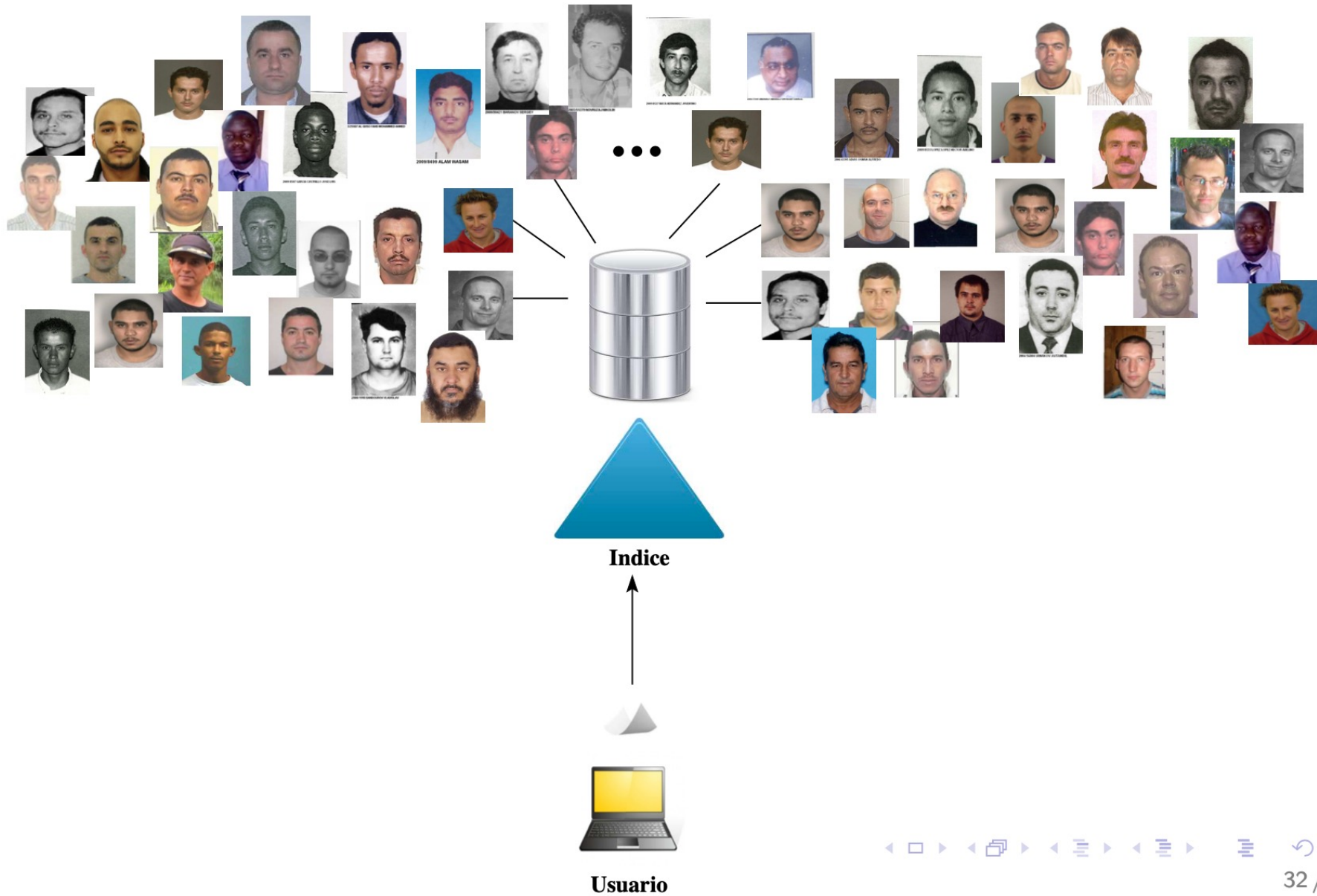
# Trabajos Publicados

# Búsqueda por Similitud

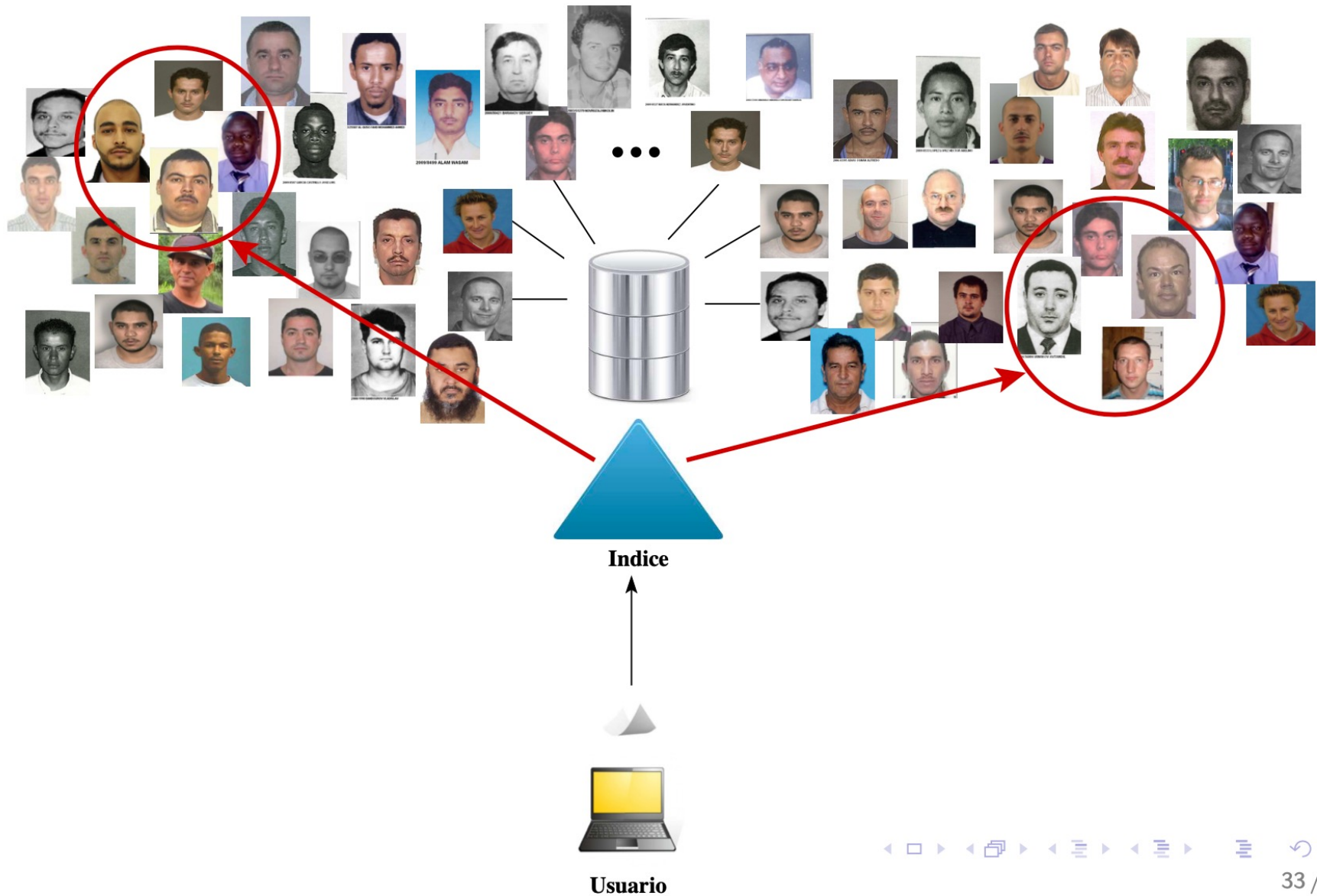




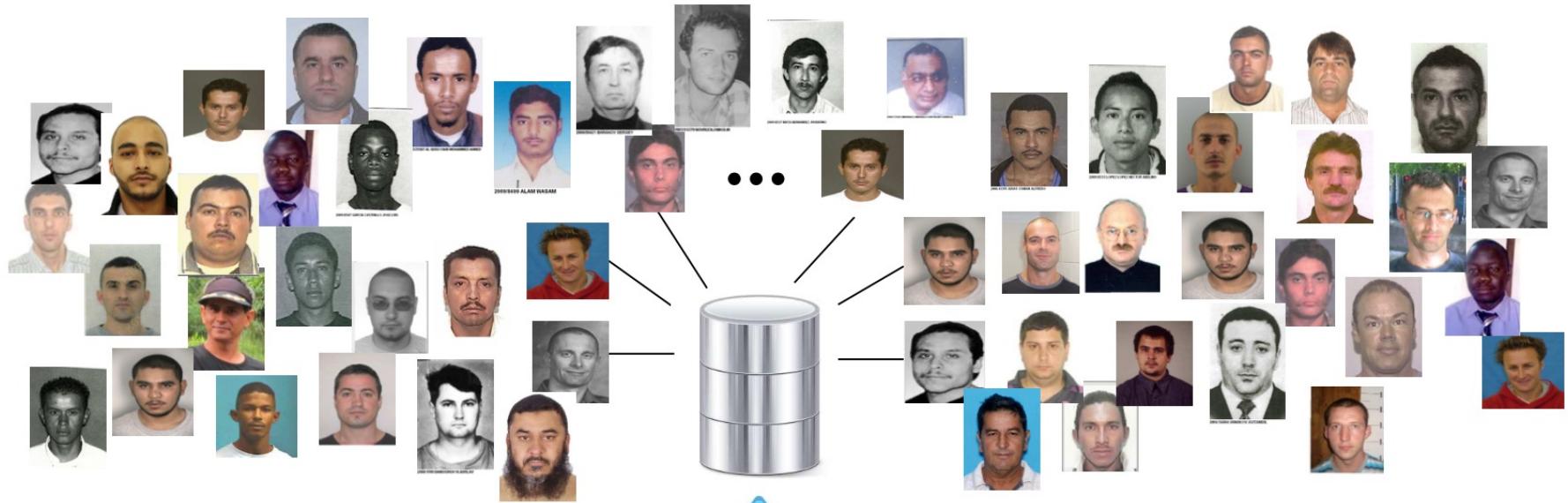
# Búsqueda por Similitud



# Búsqueda por Similitud



# Búsqueda por Similitud



**Indice**

**Objetos  
Consulta**



**Usuario**



**Usuario**

...



**Usuario**

...



**Usuario**

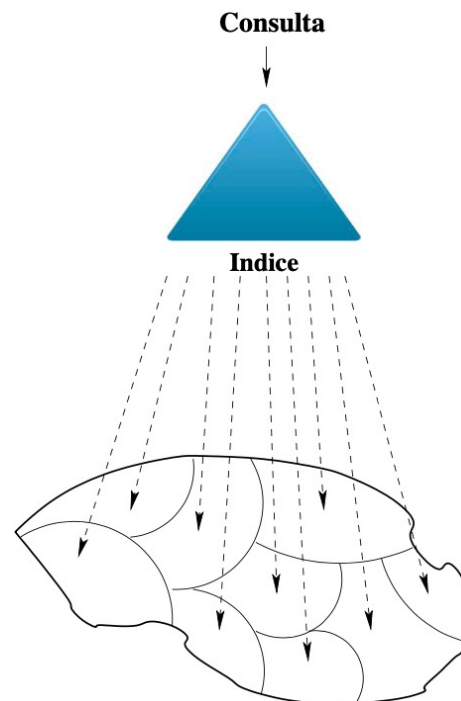


**Usuario**

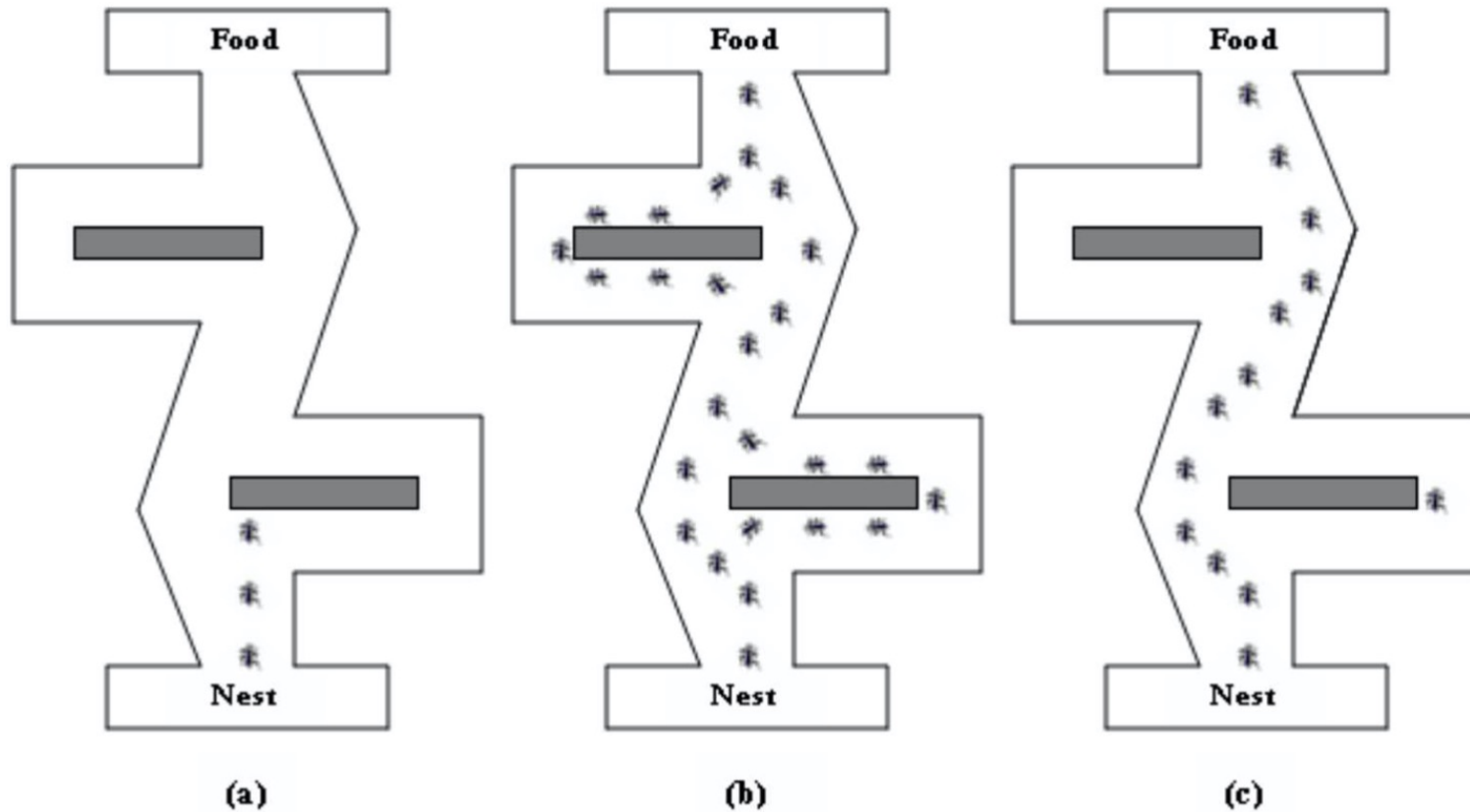


# Búsqueda por Similitud

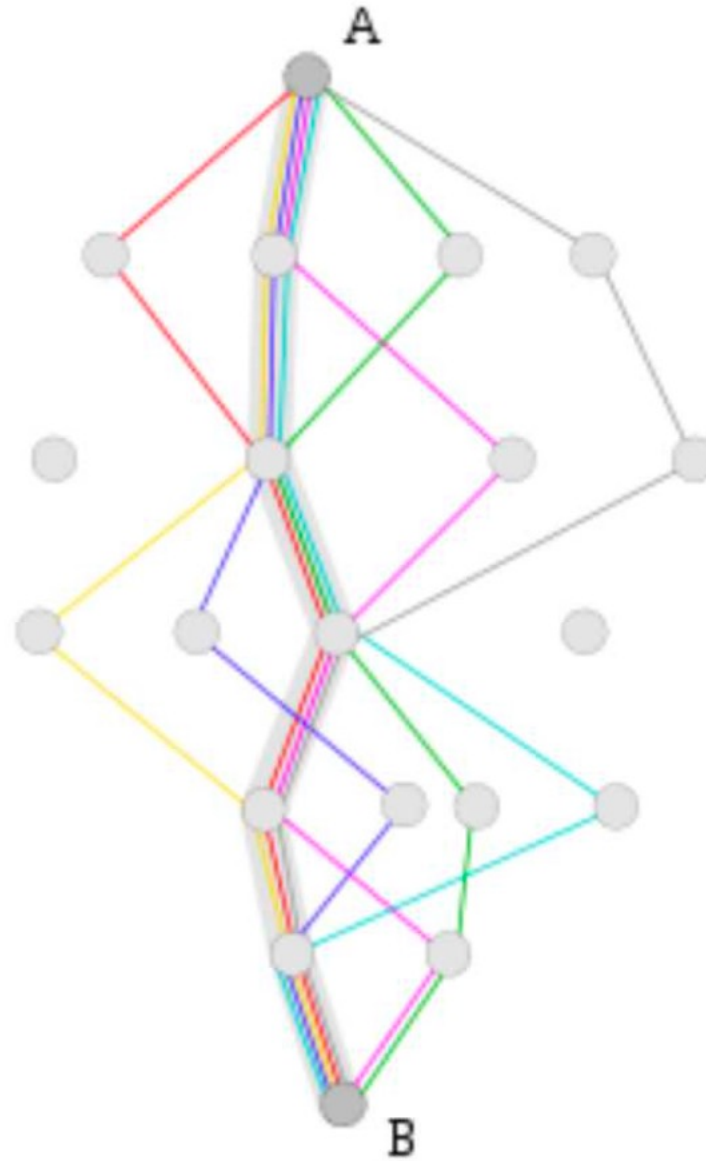
- Evaluaciones de distancia son costosas.
- Proceso off-line de la base de datos.
- Algunas distancias claves son almacenadas.
- Descarte de elementos mediante desigualdad triangular.
- Índices utilizados: EGNAT, M-Tree, SSS-Tree, SSS-Index, List of Clusters (LC).



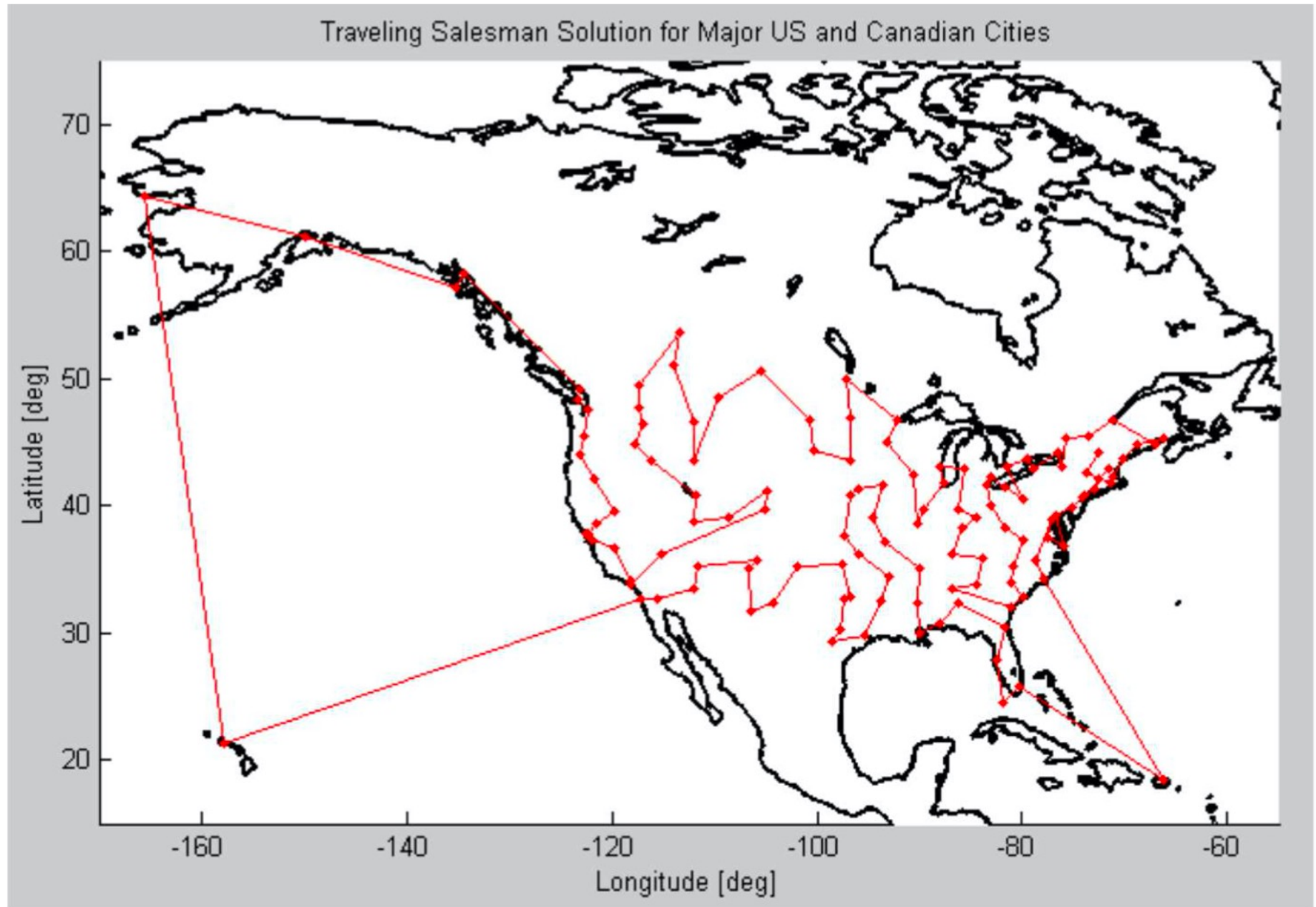
# Metaheurística: Ant Colony Optimization (ACO)



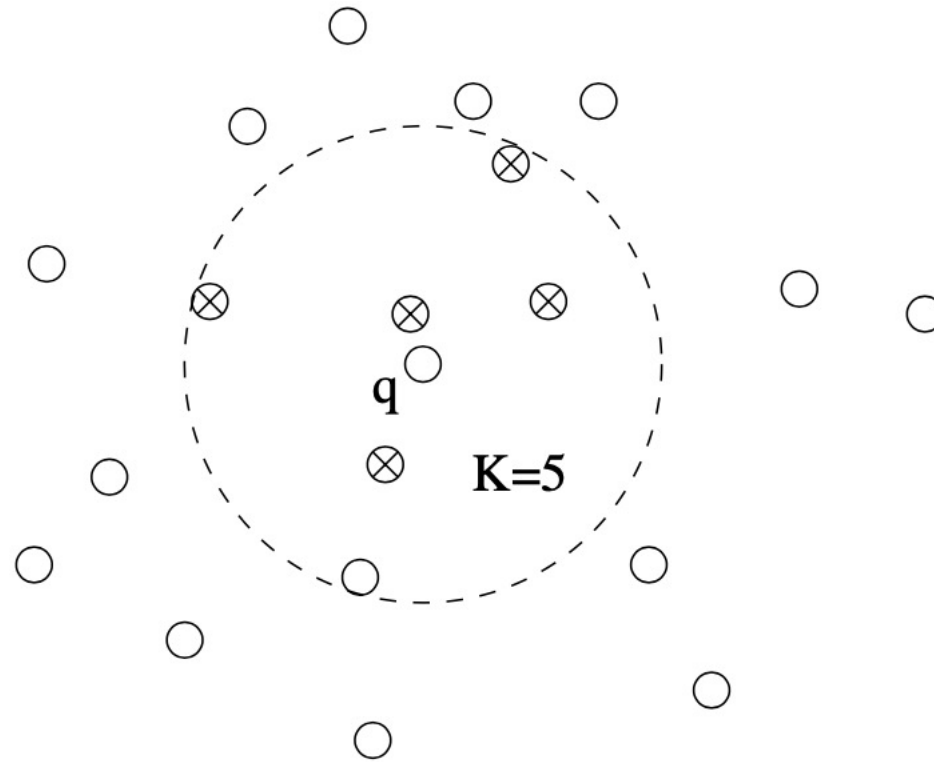
# Metaheurística: Ant Colony Optimization (ACO)



# TSP: Traveling Salesman Problem



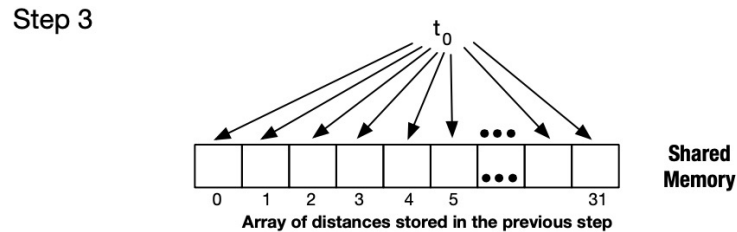
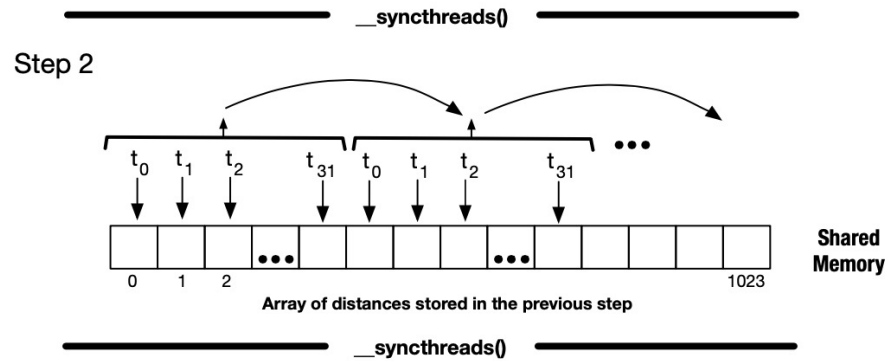
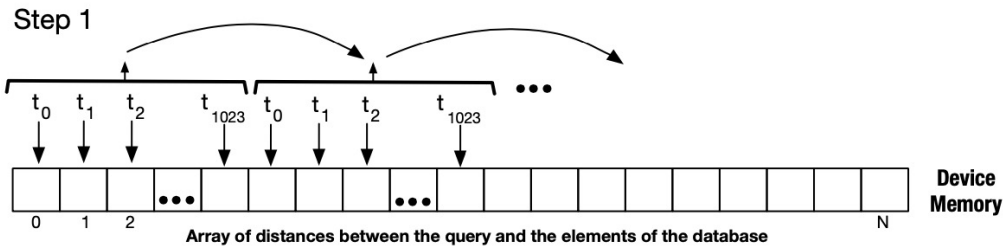
# kNN (k nearest neighbors)



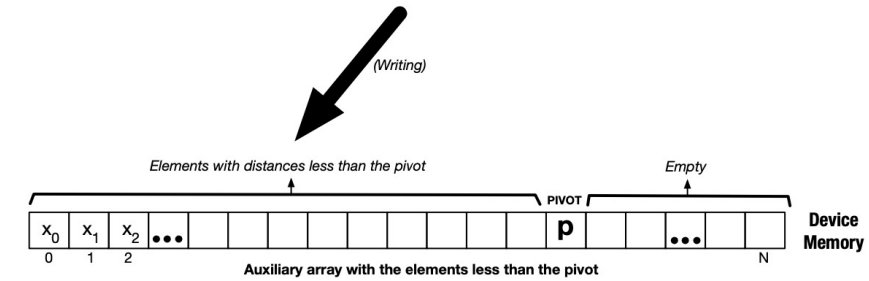
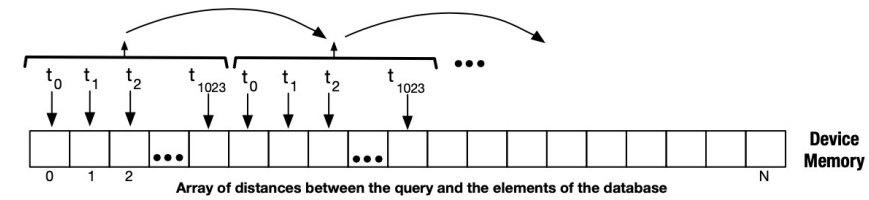


# kNN (k nearest neighbors)

## Algoritmo en GPU basado en Selection Sort



## Algoritmo en GPU basado en Quicksort



# Green Computing



## “Identificación masiva de Personas mediante Huella Digital”

Proyecto FONDEF IDeA ID17i10254  
CONICYT, Gobierno de Chile  
2018-2020



**Ricardo Barrientos<sup>(1)</sup>, Marco Mora<sup>(1)</sup>, Paulo González<sup>(1)</sup>, Wladimir Soto<sup>(1)</sup>**

Asesores Externos: **Emilio Luque<sup>(2)</sup>, Daniel Peralta<sup>(3)</sup>**

(1) Universidad Católica del Maule, Talca, Chile.

(2) Universidad Autónoma de Barcelona, España

(3) Instituto VIB, Gante, Bélgica

**Identificación Masiva de Personas Mediante Huella Digital**

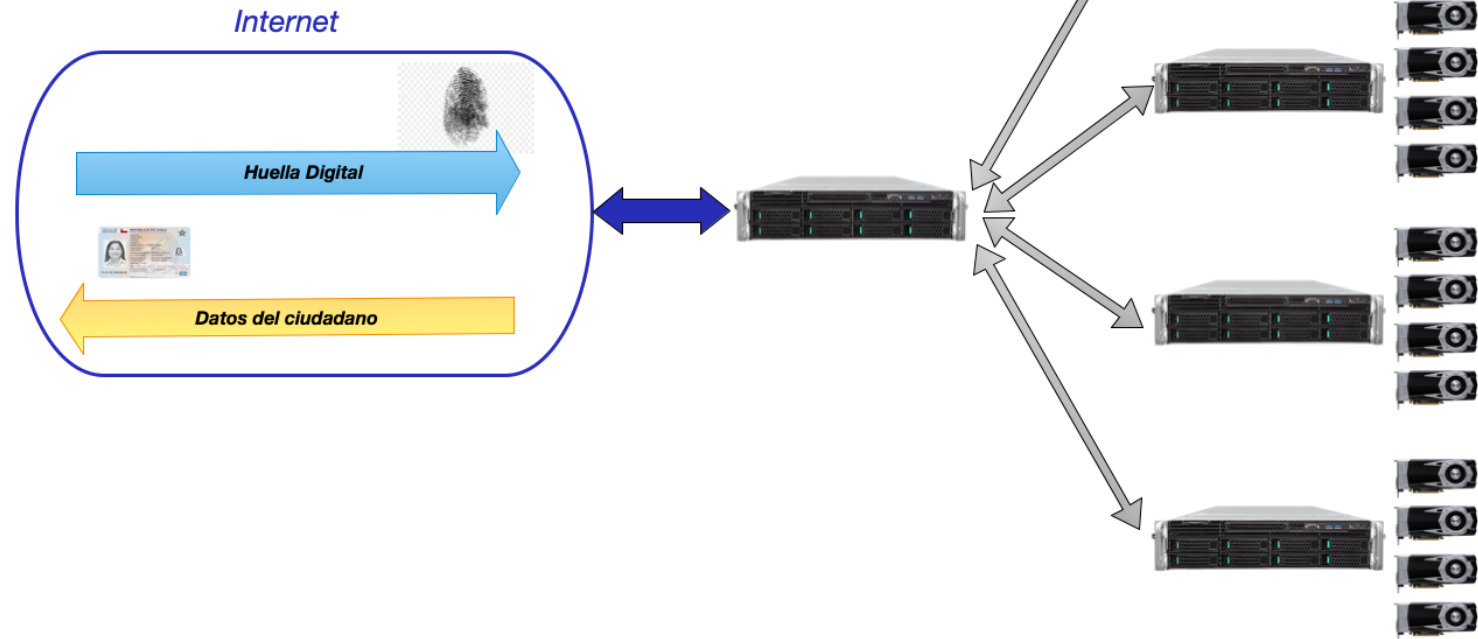


Proyecto FONDEF IDeA ID17i10254  
CONICYT, Gobierno de Chile  
2018-2019

(Opción Fija)



(Opción Móvil)



Algoritmo jerárquico de 2 etapas

# Posibles usos

## Control de Identidad Policial



## Identificación de personas en desastres naturales



## Entidades privadas

- Bancos
- Isapres
- AFP

## Identificación de personas con delitos violentos en la entrada de los estadios



## Conocer identidad de personas en caso de accidentes

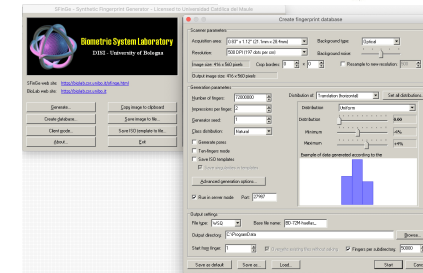


## Conocer identidad de personas en condición de calle



## Pasos de Desarrollo

1. Adquisición, Instalación y Configuración del clúster de servidores
2. Creación de la Base de Datos de huellas
3. Implementación del algoritmo en GPU
4. Implementación del algoritmo multi-núcleo
5. Desarrollo del dispositivo móvil
6. Integración de los elementos anteriores

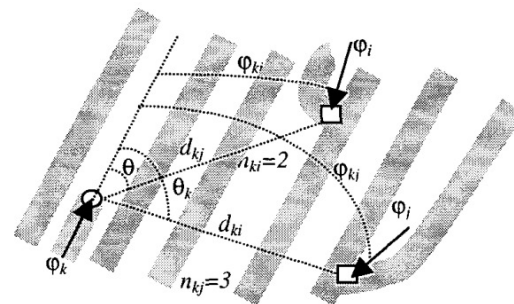
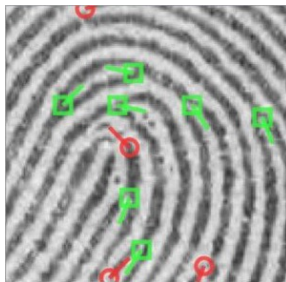


Carcasa 3D (en su interior está la tarjeta UD00)

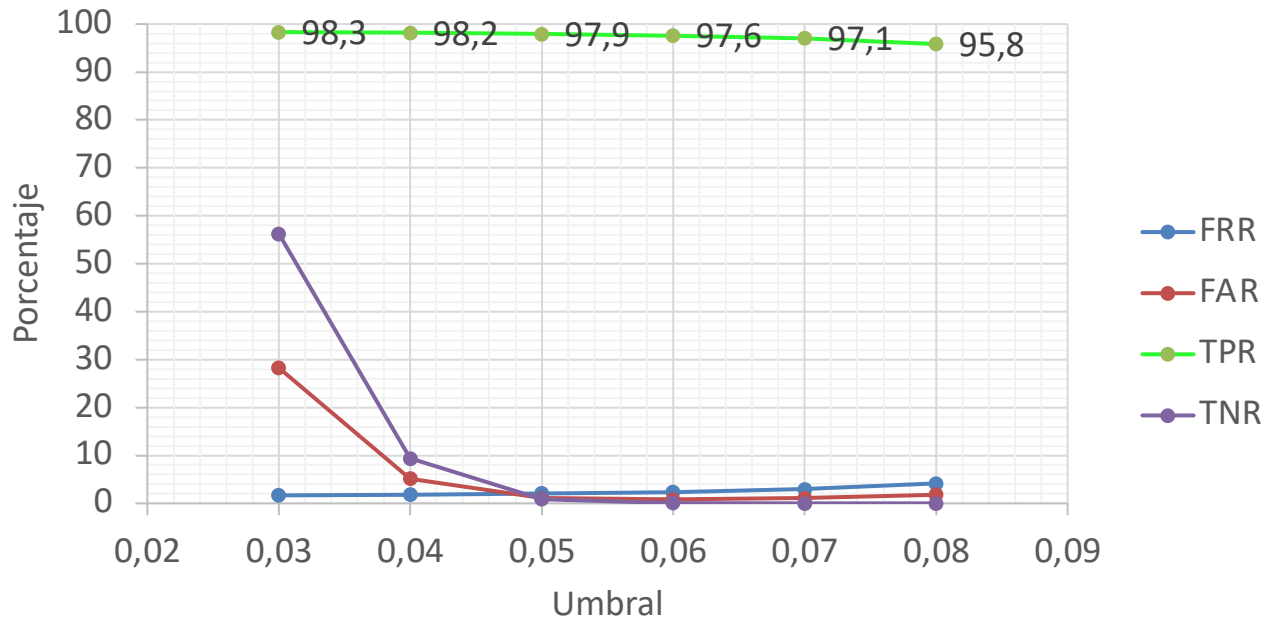
Pantalla Táctil

Batería

Sensor de captura de huella digital



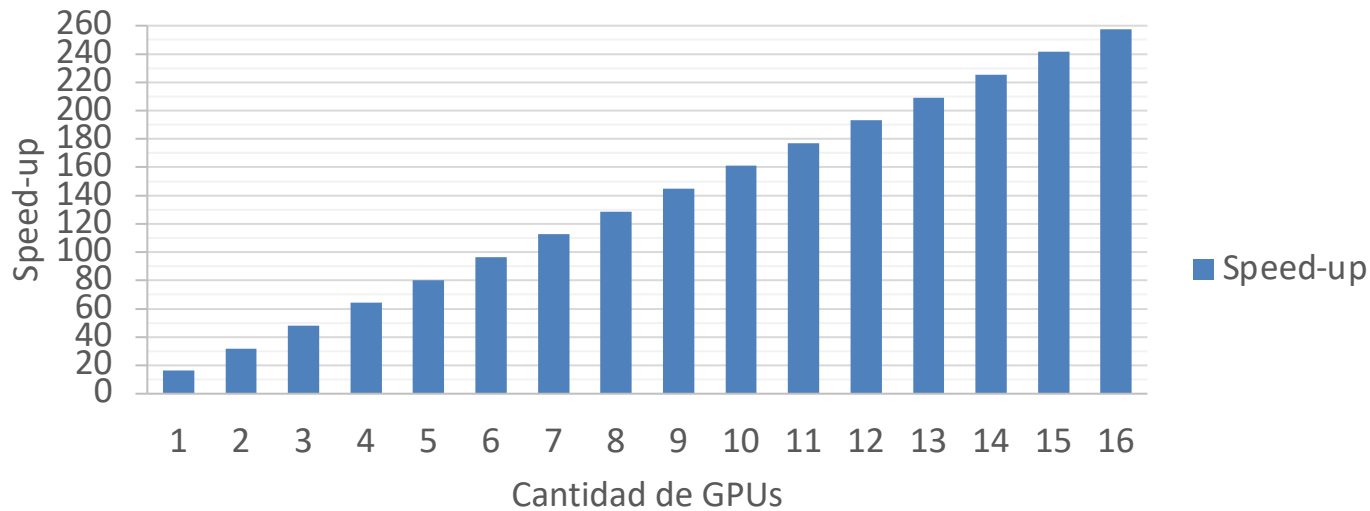
# Sistema Integrado



Valores FRR (*False Rejection Rate*), FAR (*False Acceptance Rate*), TPR (*True Positive Rate*) y TNR (*True Negative Rate*).

- Base de datos utilizada: 18.000.000 de huellas digitales
- Porcentaje de Acierto alcanzado: **97,9%**

# Sistema de Identificación



Speed-up (tiempo secuencial/tiempo paralelo): Indica cuantas veces mejor es nuestro sistema comparado con el sistema ejecutado en un procesador.

- Base de datos utilizada:  
18.000.000 de huellas digitales
- Nuestro sistema es 257 veces mejor que el programa convencional (ejecutado en 1 procesador)
- Tiempo para resolver una consulta **sin** nuestro sistema:  
6 horas
- Tiempo de nuestro sistema en resolver una consulta:  
**5,97 segundos**



## Objetivo Científico: Clasificación de una huella digital

Dado que las huellas dactilares pueden ser clasificadas en tipos según su forma (Fig.3), entrenaremos una red neuronal (Fig.4) con millones de datos de entrenamiento. De esta forma, la red neuronal será capaz de decidir, de qué tipo es una huella dada como entrada.

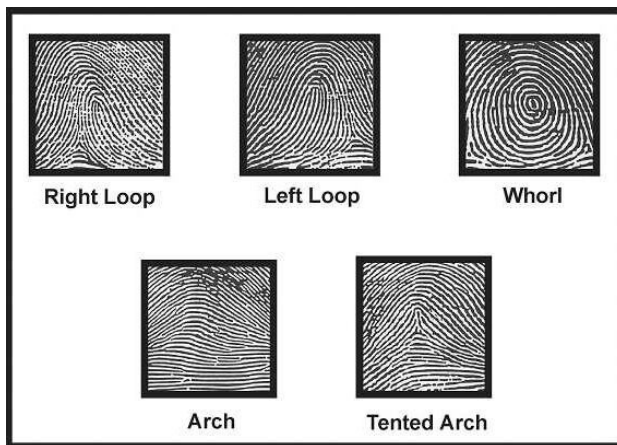


Fig.3: Tipos de huellas.

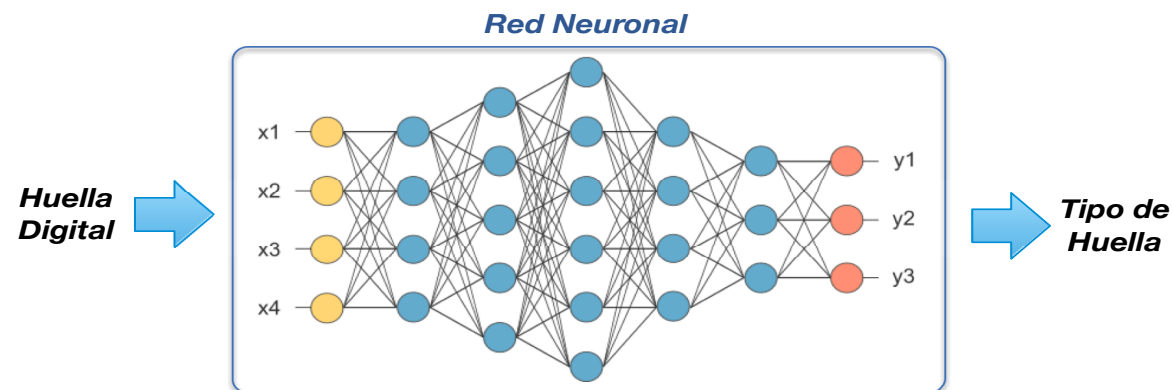


Fig.4: Red neuronal de clasificación de huellas.

## Rasgos Biométricos



Figura 2.1: Rasgos biométricos humanos



### Ventajas:

- Fácil de capturar
- Amplio desarrollo de algoritmos
- Disponibilidad de lectores

### Desventajas:

- El roce las destruye
- El envejecimiento impide su lectura
- Se puede falsificar

# Métodos Biométricos (Iris)

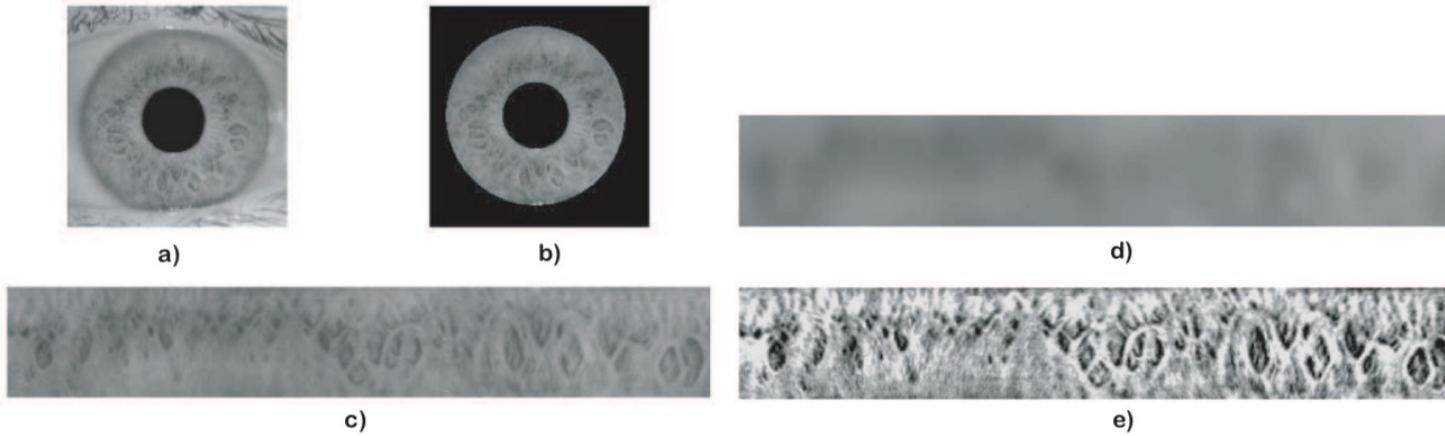


Figura 4.25: a) Imagen original, b) Iris localizado, c) Imagen normalizada, d) Iluminación de fondo estimada y e) Imagen normalizada después de la mejora de contraste.

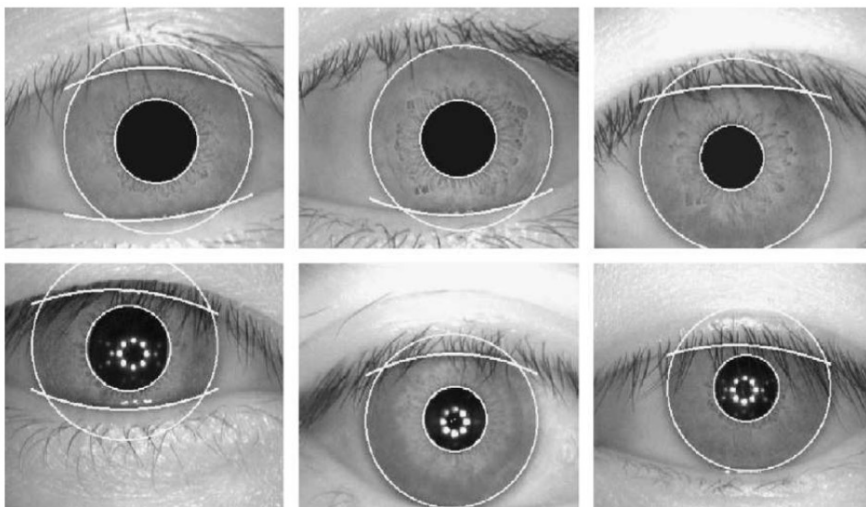
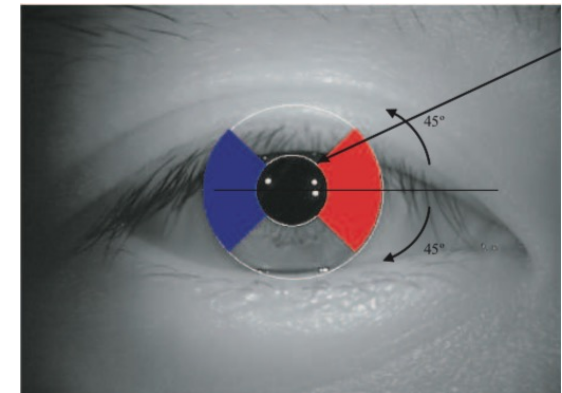


Fig.: Lectura parcial del iris.



- Ventajas y desventajas
- Utilizado por gobiernos como India

# Objetivo Científico: Identificación de personas mediante nuevos métodos biométricos.



Fig.5: Lector de venas del dedo.

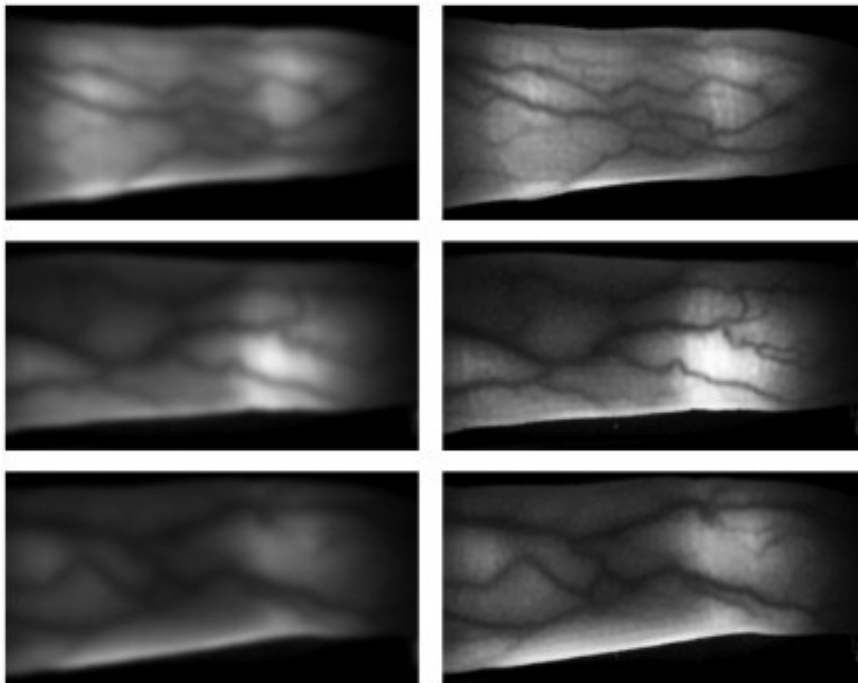


Fig.6: Ejemplo de venas del dedo.

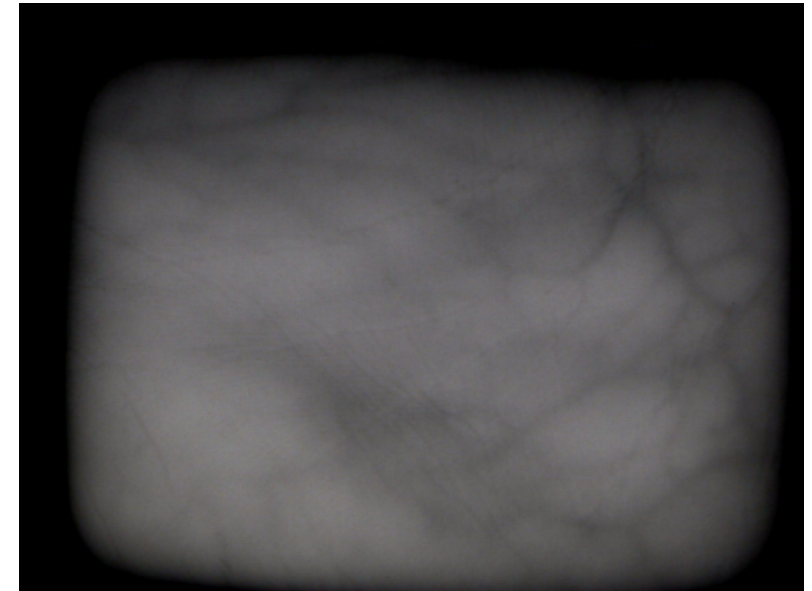


Fig.7: Ejemplo de venas de palma.

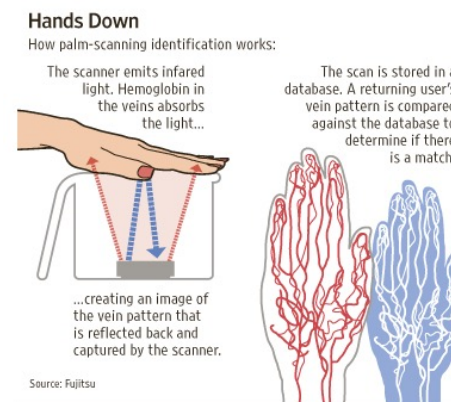


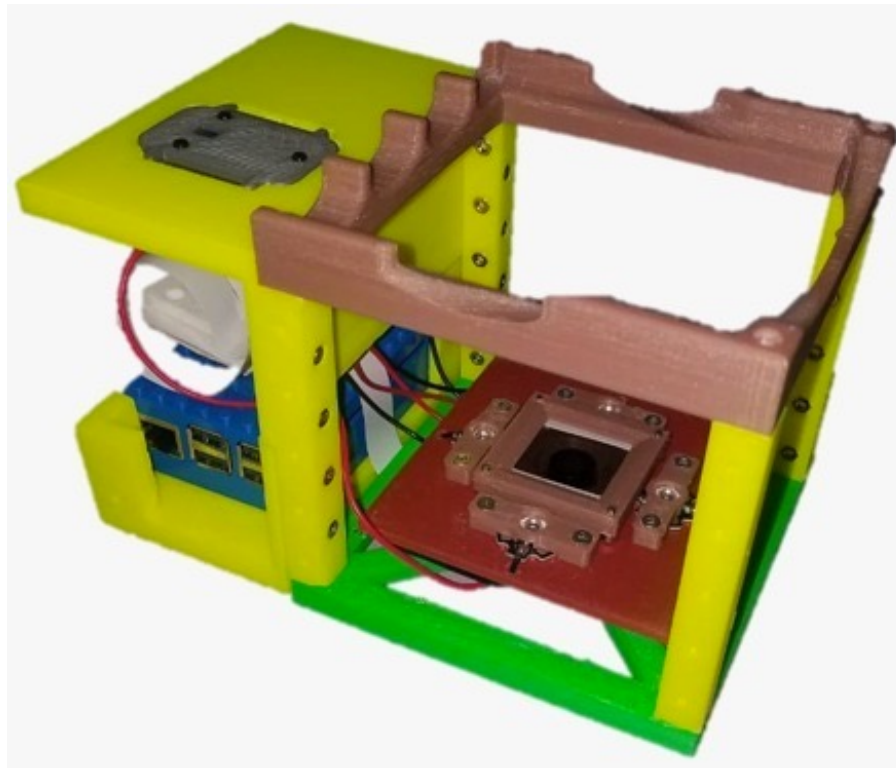
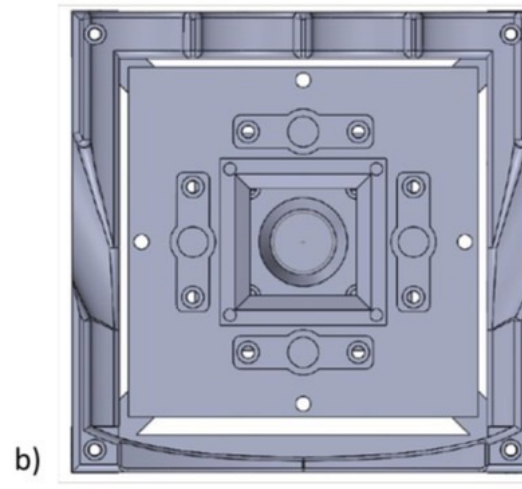
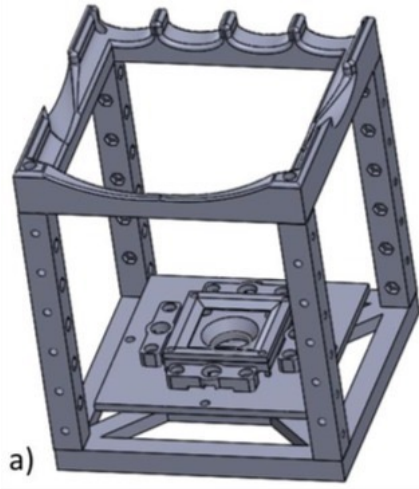
Fig.8: Lector de venas de palma.

Ventajas y desventajas

- Imagen privada y no accesible



- Prototipo



**Muchas gracias por su atención**