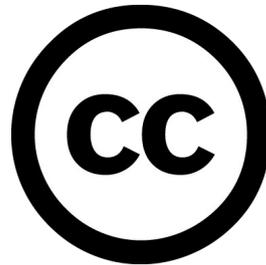


atmantree.com

El autor del presente documento lo ha publicado bajo las condiciones que especifica la licencia



Creative Commons
Attribution-NonCommercial-ShareAlike 3.0

<http://creativecommons.org/licenses/by-nc-sa/3.0/>

En caso de dudas escriba a:
info@atmantree.com

¿POR QUÉ HACEMOS
PERSISTENCIA DE DATOS?

Y PARA QUÉ

¿POR QUÉ HACEMOS
PERSISTENCIA DE DATOS?

Ing. Carlos Gustavo Ruiz
@atmantree
<http://atmantree.com>

Agenda

- Breve presentación
- Qué se entiende como persistencia de datos
- Cómo podemos hacer persistencia de datos
- Gestores de Bases de Datos
- Entre Ácidos y Bases
- Define Calidad de Servicio
- De lo micro a lo macro (embebidos -> BigData)

Presentación

Mi nombre es Carlos Gustavo Ruiz

Ingeniero en Sistemas

Consultor && Programador && Voluntario

Presentación

Mi nombre es Carlos Gustavo Ruiz

Ingeniero en Sistemas

Consultor && Programador && Voluntario

Actualmente:

CNTI
INDENE - USB
GDC - UCV
Min Defensa

...



Actualmente:

Python
C++
JavaScript
PL/pgSQL

...



Actualmente:

Python Venezuela
TuBaseDeDatosLibre.org
OpenVE
Geomática Libre Vzla

...

Persistencia de Datos

Cuando queremos saber o recordar:

Persistencia de Datos

Quando queremos saber o recordar:



Persistencia de Datos

Cuando queremos saber o recordar:



...buscamos recursos y fuentes como libros, periódicos, retratos, fotos, partituras, registros, etc.

Persistencia de Datos

Cuando queremos recordar el valor de una variable luego de la finalizar la ejecución de un programa entonces..

Persistencia de Datos

Cuando queremos recordar el valor de una variable luego de la finalizar la ejecución de un programa entonces..

Hacemos persistencia de datos

El cómo de la Persistencia de Datos

Desde el punto de vista del Hardware podemos ver a la persistencia como una “evolución” de medios de almacenamiento

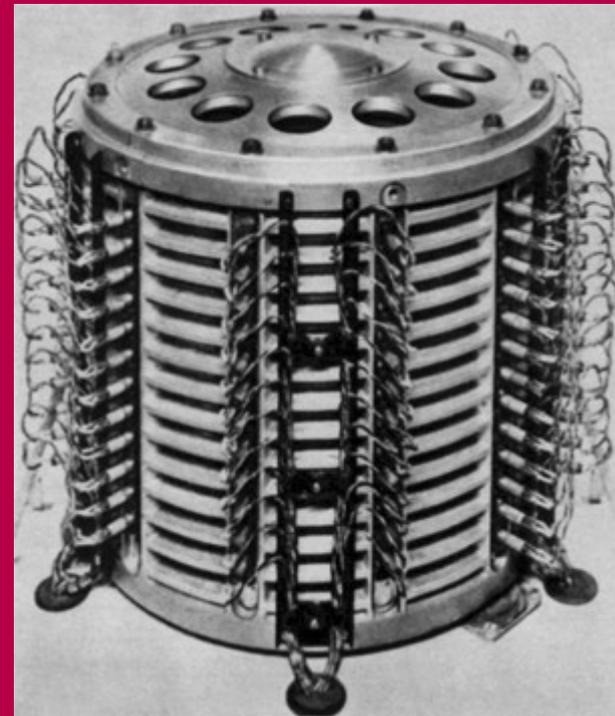
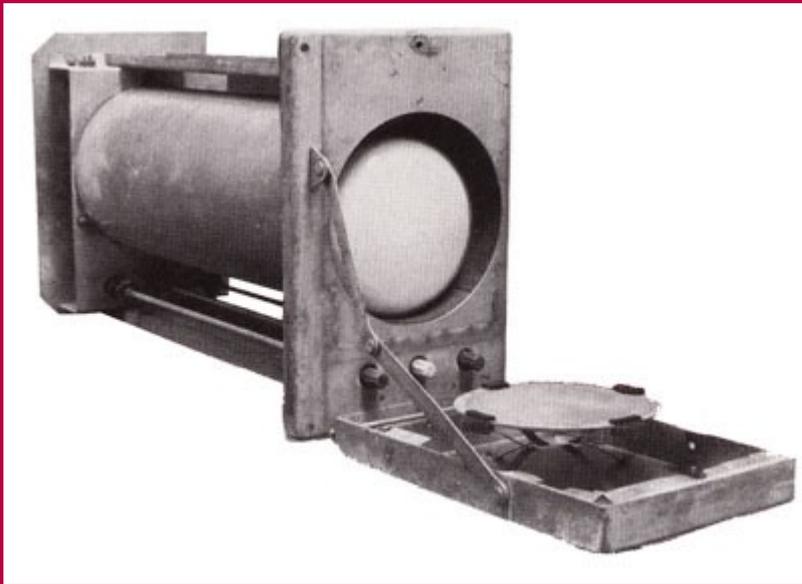
El cómo de la Persistencia de Datos

Tarjetas Perforadas (1725-1880)



El cómo de la Persistencia de Datos

Tubo de Williams (1945) y Memoria de Tambor (1950)



El cómo de la Persistencia de Datos

Cintas Magnéticas (1951) y Discos Duros (1956)



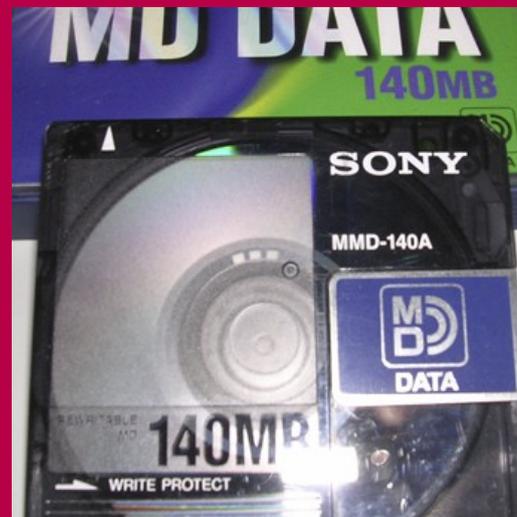
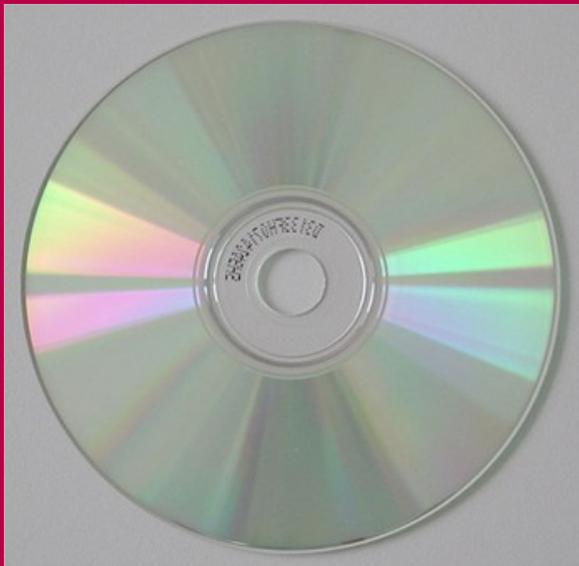
El cómo de la Persistencia de Datos

Etapa de la Miniaturización (1972 - 1987)



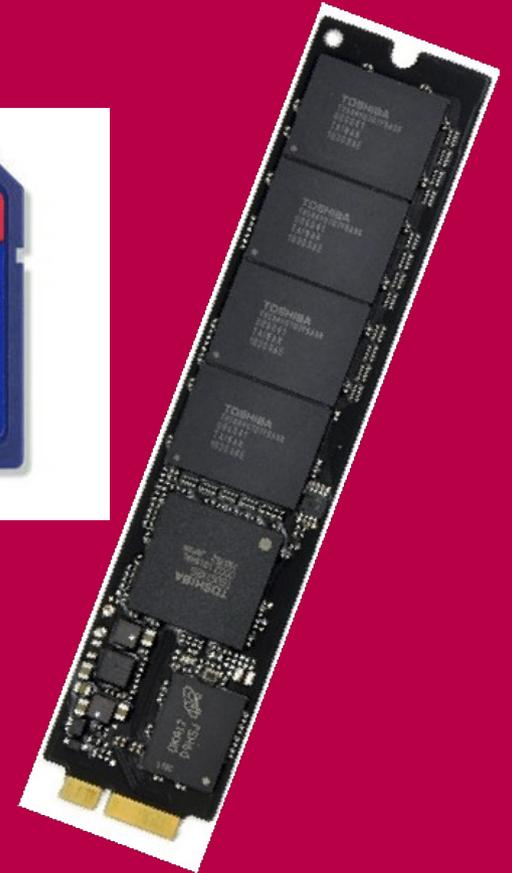
El cómo de la Persistencia de Datos

La Etapa del Laser (1990 - 2008)



El cómo de la Persistencia de Datos

La Nueva Etapa del Almacenamiento



El cómo de la Persistencia de Datos

Desde el punto de lógico se ha avanzado en mejorar el almacenamiento de datos a partir de 2 tipos de archivos básicos, textos planos y binarios.

Sin embargo la potencia no viene estrictamente de la estructura básica del medio de almacenamiento sino de los mecanismos para hacerlo.

El cómo de la Persistencia de Datos

Ejemplos de archivos de datos en texto plano:

- Texto delimitado
 - Por columnas, comas, caracteres especiales o tabuladores
- Lenguajes marcados
 - El más difundido es XML
- Formatos de Intercambio
 - El favorito del momento es JSON

El cómo de la Persistencia de Datos

Ejemplos de archivos de datos en formato binario:

- Generalmente cualquier dato almacenado en una suite ofimática está en formato binario
- Tradicionalmente se le colocaba la extensión .dat a este tipo de archivos
- Información de niveles digitales como señales o píxeles (en caso de imágenes)
- Archivos de texto plano comprimidos

El cómo de la Persistencia de Datos

Para estructurar los datos entonces se recurren a diversos métodos:

- Organización Jerárquica de acceso al filesystem (siguiendo la estructura de directorios)
- Manejados desde una capa gestora de abstracción de los datos (separando la implementación de la presentación)

Gestores de Bases de Datos

Los Gestores de bases de datos son una capa de abstracción que permite desentendernos de los pormenores de guardar el dato y concentrarnos en lo que para nosotros es importante, la lógica de nuestros procesos y del negocio.

Un buen gestor de base de datos debe garantizar la durabilidad de los datos, sin embargo también debe parecerse a nuestro trabajo.

Gestores de Bases de Datos

En tal sentido, no queda duda que CAP es el origen de los problemas al elegir un gestor de base de datos.

Gestores de Bases de Datos

Consistency
Availability
Partition Tolerance

Gestores de Bases de Datos

Consistency: que todos los nodos vean la misma información al mismo tiempo

Gestores de Bases de Datos

Availability: la garantía de que cada petición a un nodo reciba una confirmación de si ha sido o no resuelta satisfactoriamente

Gestores de Bases de Datos

Partition Tolerance: que el sistema siga funcionando a pesar de algunas pérdidas arbitrarias de información o fallos parciales del sistema a menos que todos los nodos fallen el sistema puede seguir funcionando

Gestores de Bases de Datos

¿Qué hay con esto?

Gestores de Bases de Datos

Que solo puedes tomar
simultáneamente 2 de los 3

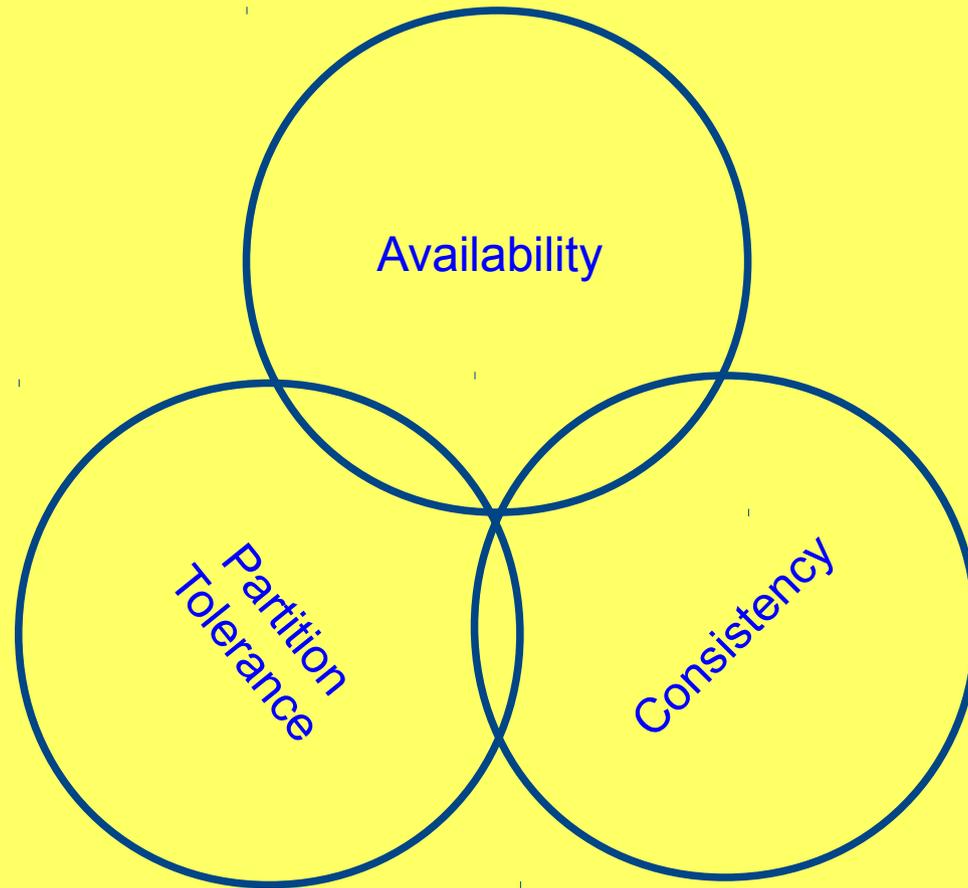
Gestores de Bases de Datos

Si, como cuando eliges entre:
tener vida social,
buenas calificaciones y
dormir

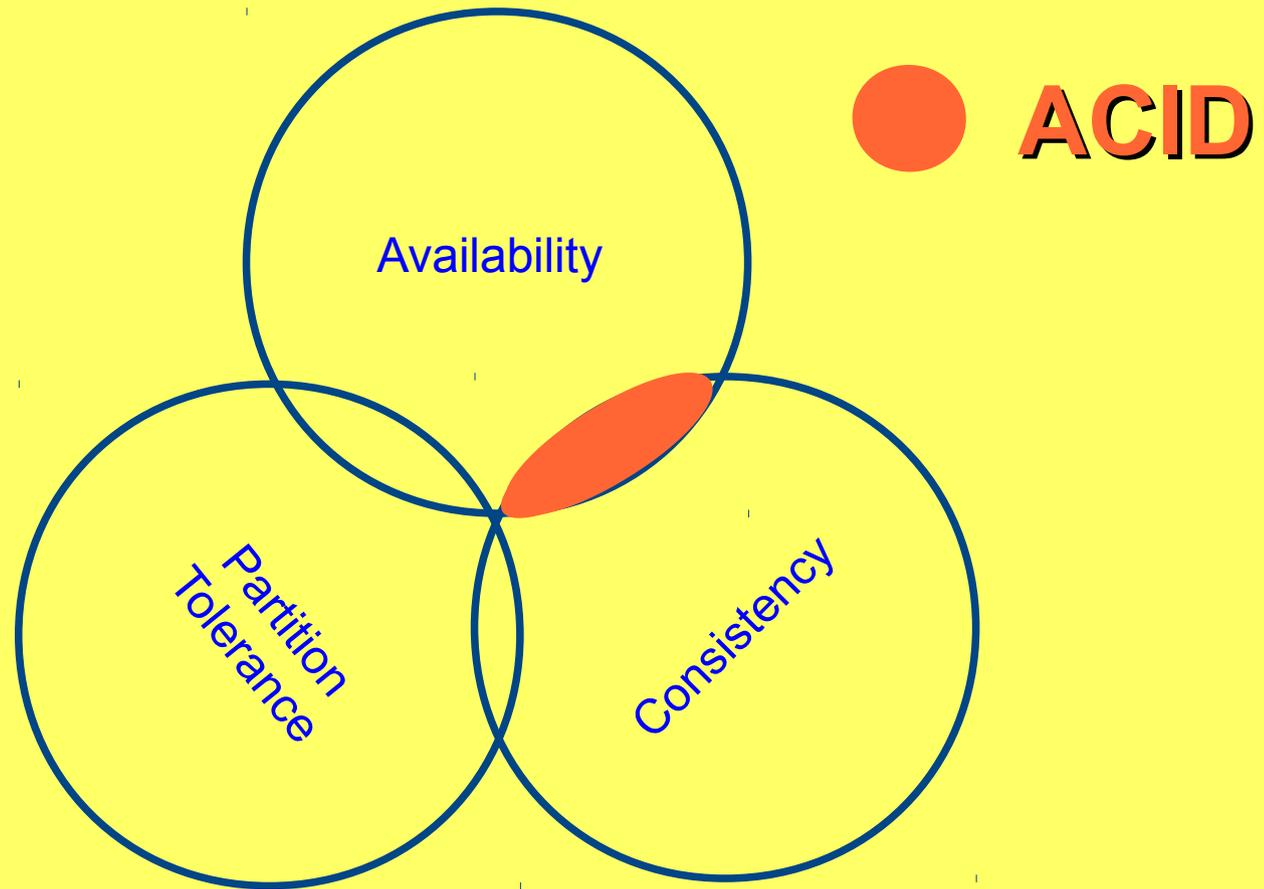
Ácidos y Bases

Si solamente podemos seleccionar 2 de las tres características dentro de un sistema de cómputo distribuido entonces hay que hacer una selección adecuada.

Ácidos y Bases



Ácidos y Bases



Ácidos y Bases

Atomicity
Consistency
Isolation
Durability

Ácidos y Bases

Atomicity: propiedad que asegura que la operación se ha realizado o no, y por lo tanto ante un fallo del sistema no puede quedar a medias. (Atomicidad)

Ácidos y Bases

Consistency: propiedad que asegura que sólo se empieza aquello que se puede acabar. Por lo tanto se ejecutan aquellas operaciones que no van a romper las reglas y directrices de integridad de la base de datos

Ácidos y Bases

Isolation: propiedad que asegura que una operación no puede afectar a otras. Esto asegura que la realización de dos transacciones sobre la misma información sean independientes y no generen ningún tipo de error.
(Aislamiento)

Ácidos y Bases

Durability: propiedad que asegura que una vez realizada la operación, ésta persistirá y no se podrá deshacer aunque falle el sistema. (Durabilidad)

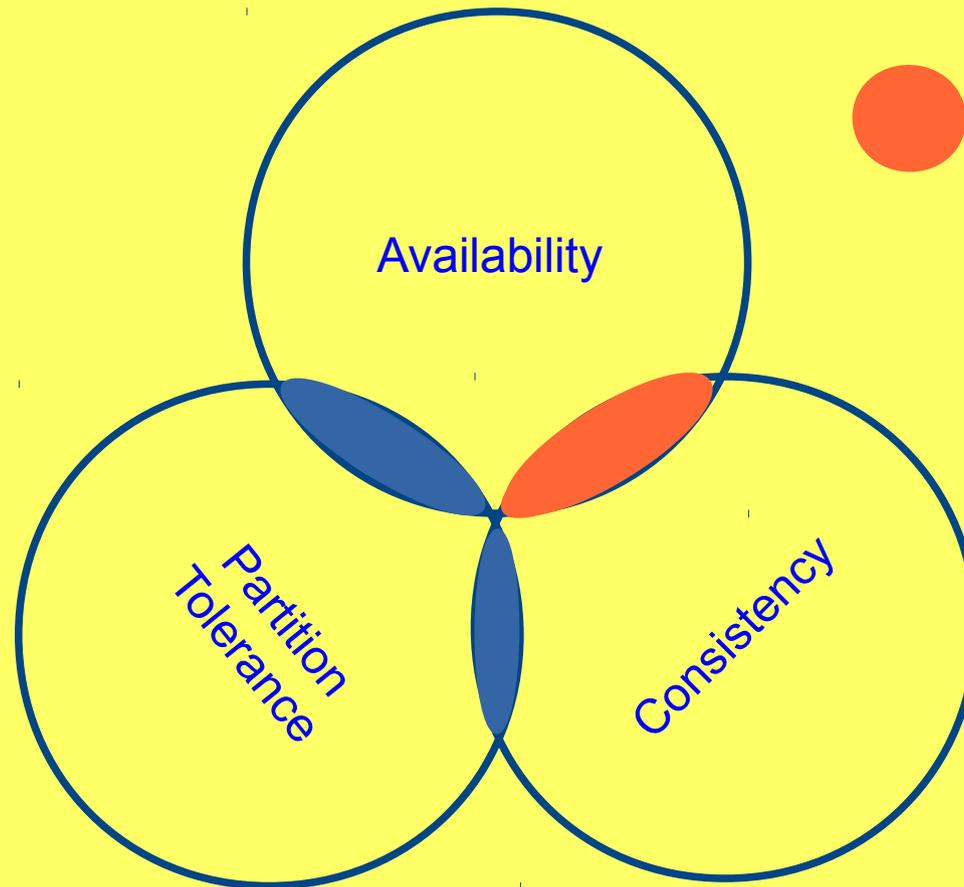
Ácidos y Bases

En este renglón se encuentran las bases de datos relacionales y SQL como idioma casi universal. Es un estándar para muchos aspectos de la industria.

Ácidos y Bases

● **BASE**

● **ACID**



Ácidos y Bases

Basically Available
Soft state
Eventual consistency

Ácidos y Bases

BASE: utiliza un enfoque optimista permitiendo que la consistencia se entienda como un estado de cambio continuo. Este enfoque es ideal para manejo de gran volumen de datos.

Ácidos y Bases

Este enfoque agrupa a las bases de datos que se conocen como NoSQL. Es un nuevo estándar para las aplicaciones en la internet pues permite escalar horizontalmente de forma transparente.

Calidad

Antes de comenzar a hablar de la calidad, vamos a hacernos de nuevo la siguiente pregunta.

Calidad

Antes de comenzar a hablar de la calidad, vamos a hacernos de nuevo la siguiente pregunta.

¿por qué hacemos
persistencia de datos?

Calidad

La primera respuesta pudiese ser que es porque los programas terminan su ejecución y no quisiera tener que reescribir todas mis cartas cada vez que enciendo el PC.

Calidad

Sin embargo esta es una respuesta ingenua.

Calidad

Persistimos los datos porque tenemos procesos asociados a esos programas que inician y terminan su ejecución.

Nóminas, Mensajería, Ofimática, Sistemas de Análisis, Modelado y Simulaciones, Dibujo y Diseño Asistido por Computador, CRM, ERP, Prospección Petrolera, Procesos Industriales, Asistentes de Vuelo, Encuestas, Planificación, Promoción, Planes de Formación.. y un largo etcétera.

Calidad

Sin esos datos almacenados y disponibles luego de finalizar nuestros programas entonces nos encontraremos sin las referencias para recordar que pasó ayer.

Calidad



Calidad

La calidad dependerá de nuestra siguiente pregunta..

Calidad

La calidad dependerá de nuestra siguiente pregunta..

¿para qué hacemos
persistencia de datos?

Calidad

Dependiendo del uso que vaya a darse los datos se tendrá que dar prioridad a alguna de las tres características de los sistemas distribuidos.

Calidad

Consistency

Availability

Partition Tolerance

Calidad

La calidad viene entonces dada por el uso, si su proceso está asociado a una nómina, por ejemplo la consistencia eventual es inaceptable.

Calidad

En cambio si usted se encarga de indexar la vida social de sus usuarios, el parloteo de millones de personas en la red, o el contenido de las páginas de la internet, entonces el tema de la consistencia eventual se convierte en LA alternativa debido a la dimensión de su problema.

..de lo micro a lo macro

Veamos un ejemplo cotidiano.. al menos hoy en día..

¿Alguien tiene un celular
como alguno de estos?

..de lo micro a lo macro



..de lo micro a lo macro

Bueno tal vez no esos modelos exactamente, pero hay buen chance que tengan un smartphone.

Bueno, su teléfono es un buen ejemplo de las bases de datos y sus enfoques actuales.

..de lo micro a lo macro

Si su amigo Juan (todos tenemos un amigo llamado Juan) escribe una actualización en Facebook, ¿es realmente importante que la reciba al segundo de publicada?

..de lo micro a lo macro

Probablemente lo importante es que eventualmente usted reciba el mensaje en su timeline y se entere que Juan es papá.

Unos minutos no harán gran diferencia.

..de lo micro a lo macro



..de lo micro a lo macro

Sin embargo, en nuestro celular no es aceptable que eventualmente podamos realizar una llamada. Es por ello que especialmente los sistemas prepago deben asegurar saber al instante cual es su saldo.

..de lo micro a lo macro

Por otro lado dentro de su dispositivo hay pequeñas bases de datos como la que contiene su lista de contactos. Puede que eventualmente concilie la información con un servicio en red, sin embargo debe estar disponible inmediatamente al momento de buscar en la copia local.

Algo más?

Algo más?

Preguntas
Comentarios
Dudas

Algo más?

Información de Contacto

@atmantree

<http://atmantree.com>

<http://tubasededatoslibre.org>