

Nicolas Arturo Alvarado 201630444

Javier David Peniche Calderón 201716772

Luimarco Daniel Santiago Carrascal Diaz 201620630

Parte A: Entendimiento del problema

Personas que usaran el sistema: Las unidades inteligentes que se instalan en las unidades de transporte, Los administradores del sistema, Los clientes (tanto empresas como personas), Los empleados de la empresa.

Transacciones que harán los usuarios de la plataforma: Envío de datos de geolocalización, Consultas a los sistemas de rastreo, Consultas al ERP y CRM, Creación y consulta de facturas y envíos.

1. Carga de trabajo

Tipo de transacción	Quien?	Que hace?	Como lo hace?	Carga
Envío de datos GPS	Unidades inteligentes	Envía su posición al servidor.	Por medio de la red celular se conecta al servicio WS-SOA.	Tamaño de la información enviada.
Consultas al sistema de rastreo	Administradores	Maneja la información relacionada con los puntos de recogida y entrega de paquetes. También maneja las rutas de las unidades de distribución.	Se conecta al servicio por medio del portal web	Tamaño de la información consultada y/o modificada.
Consultas al sistema de rastreo	Clientes	Rastrea los paquetes que pidió o envió.	Se conecta al servicio por medio del portal web	Tamaño de la información consultada.
Consultas ERP	Administradores	Maneja y consulta la información de los recursos de la empresa	Se conecta al servicio por medio del portal web	Tamaño de la información consultada.
Consultas CRM	Administradores	Maneja y consulta la información de los clientes.	Se conecta al servicio por medio del portal web	Tamaño de la información consultada.

Crear envío o "contrato"	Clientes	Envía la información requerida para la generación de un contrato. En el caso de los clientes naturales también se crea una factura, para las empresas la factura se crea a final de mes.	Se conecta al servicio por medio del portal web	Tamaño de la información enviada. Y en caso de crear la factura, lo anterior más el tamaño de la factura.
Consulta de contratos.	Administradores	Maneja y la consulta la información de los contratos.	Se conecta al servicio por medio del portal web	Tamaño de la información consultada.
Consulta de facturas	Administradores	Maneja y la consulta la información de las facturas.	Se conecta al servicio por medio del portal web	Tamaño de la información consultada.

2. Requerimientos por aplicación

Desempeño

Se tomarán los datos de temporada alta ya que este es el más exigente para el sistema.

Aplicación	Contrato de transporte	Consulta ERP	Consulta CRM	Facturas	Recogida de paquetes	Rastreo de paquetes	Rastreo de unidades	Total
ERP	50000	1000	0	35000	0	0	0	86000
CRM	50000	0	500	0	7000	0	0	57500
DMS	50000	0	0	0	0	15000	0	65000
GDTs	0	0	0	0	7000	0	2500	9500

Total de transacciones por día en temporada alta: 218000

Podemos ver que las aplicaciones ERP, CRM y DMS son las que más transacciones reciben por día en temporada alta, especialmente el ERP, este es el que más transacciones recibe de todos. Esto se tiene que tener en cuenta para la infraestructura ya que estas son las aplicaciones que más recursos consumirán.

También toca tener en cuenta que la plataforma de integración tiene que poder soportar el total de peticiones ya que a esta le llegan todas las peticiones.

ERP: Como se mencionó anteriormente, esta será la aplicación a la que más atención se le debe tener ya que es la que más transacciones por día recibe.

CRM: Esta aplicación recibe un número alto de transacciones por lo que toca tenerla en cuenta.

DMS: Esta aplicación recibe un número alto de transacciones por lo que toca tenerla en cuenta.

GDTs: Esta aplicación tiene un número de transacciones bajo, esto no significa que no se tenga que tener en cuenta, sino que no afecta mucho en la decisión final de la arquitectura.

Plataforma de integración: Esta es la parte más importante de todo el sistema ya que si se cae todo cae. Claro que los procesos van a ser resueltos por las aplicaciones como tal y no la plataforma. Por esto no se requiere de tanto desempeño en la plataforma a pesar de ser la que recibe todas las peticiones.

Capacidad

Información que requiere persistencia en disco: Información de las 400 unidades de transporte, Información de los envíos (info del paquete + contrato), Información de los usuarios, Las rutas de las unidades de distribución (actualmente son 100 y se espera en 4 años manejar 400) y por último las facturas.

ERP: Este requiere tener las facturas más la nómina de la empresa.

CRM: Este requiere tener toda la información de los usuarios (tanto personas como empresas).

DMS: Este requiere tener la información de las unidades de transporte y la información de los envíos.

GDTs: Este requiere tener la información de las rutas de las unidades de distribución.

Plataforma de integración: La plataforma no guarda mucha información y la poca que guarda no varía con el tiempo.

Para el almacenamiento de las aplicaciones no se tiene suficiente información para crear un plan detallado de infraestructura. De todas maneras, estas manejan información crucial para el funcionamiento de la empresa y que en los próximos 4 años se espera que aumente, por lo que se debe tener un buen sistema de almacenamiento para estas aplicaciones.

Escalabilidad

Se debe tener en cuenta que: En el año actual 100 de las 400 unidades de transporte tienen dispositivos inteligentes y se planea que el 100% tenga en 4 años (aumento de 75 unidades por año).

Este aumento impactará principalmente al GDTs ya que este es el encargado del rastreo de las unidades de distribución.

Se está planeando que la infraestructura soporte el sistema en temporada alta, por lo que el cambio entre temporada baja y alta no es relevante. El cambio que tendrá que soportar toda la infraestructura es el que resulte de adquirir nuevos clientes o que los clientes actuales hagan más pedidos. Este es un valor difícil de estimar con los datos que tenemos, pero de todas maneras toca tener en cuenta que se tiene que poder escalar el sistema para que la compañía pueda aceptar más pedidos y usuarios.

ERP, CRM, DMS: Tienen que escalar para poder aceptar más pedidos y usuarios en el futuro.

GDTs: Dado que actualmente con las 100 unidades inteligentes este servicio recibe 2500 peticiones por día y se planea aumentar las unidades inteligentes en 75 cada año. Se puede estimar que al año se verá un aumento de 1875 peticiones a este sistema. O, en otras palabras, el sistema tendrá que resistir un aumento de 25 peticiones diarias por cada unidad nueva que se adquiera.

Plataforma de integración: Esta tiene que poder soportar el aumento total de todas las aplicaciones ya que todas las peticiones pasan por esta plataforma.

Disponibilidad

Para cada aplicación se quiere tener 99.92% de disponibilidad. Para la plataforma de integración se quiere tener una disponibilidad muy alta ya que el funcionamiento de esta es indispensable para todo el sistema (se querrá una disponibilidad del 99.99% para que este no haga de cuello de botella frente a las otras aplicaciones). Se cuenta entonces con un cluster local con failover remoto para la implementación del sistema.

Para calcular la disponibilidad de todo el sistema primero identificamos la disponibilidad de las 4 aplicaciones, que están (DMS, GTDS, CRM y ERP) en paralelo con una disponibilidad del 99.92%, que estarían a la vez en serie en serie con el sistema de integración con disponibilidad del 99.99%. El cálculo se muestra en la siguiente imagen:

$$(1 - ((1 - 0.9992)^4)) \cdot 0.9999 \cdot 100 = 99.98999$$

El sistema como conjunto tiene una disponibilidad de 99.98999%, lo que corresponde a una inoperatividad en tiempo de 53 minutos y 5 segundos al año.

Parte B: Selección de la infraestructura

1. Problemática por aplicación

ERP (Escalabilidad, Capacidad, Desempeño): El ERP recibe y maneja las facturas, los contratos y la información financiera de la empresa y aunque no sabemos a qué velocidad, se asume que el número de clientes, pedidos y empleados incrementara con el tiempo. Estos tres implican un incremento en el uso del ERP por esto es importante que escale. Aunque no sabemos numéricamente cuanta capacidad requiere, se puede concluir dado que el ERP maneja la información que maneja que va a requerir una alta capacidad de almacenamiento. Esta aplicación es la que más recibe peticiones de todas y se espera que el número de peticiones incremente, por esto también es importante el desempeño.

CRM (Escalabilidad): Como se mencionó anteriormente, se espera que el número de clientes aumente con el tiempo y ya que el sistema es el encargado de la información de los clientes, este tiene que poder escalar para que no se vea afectado con la llegada de nuevos clientes.

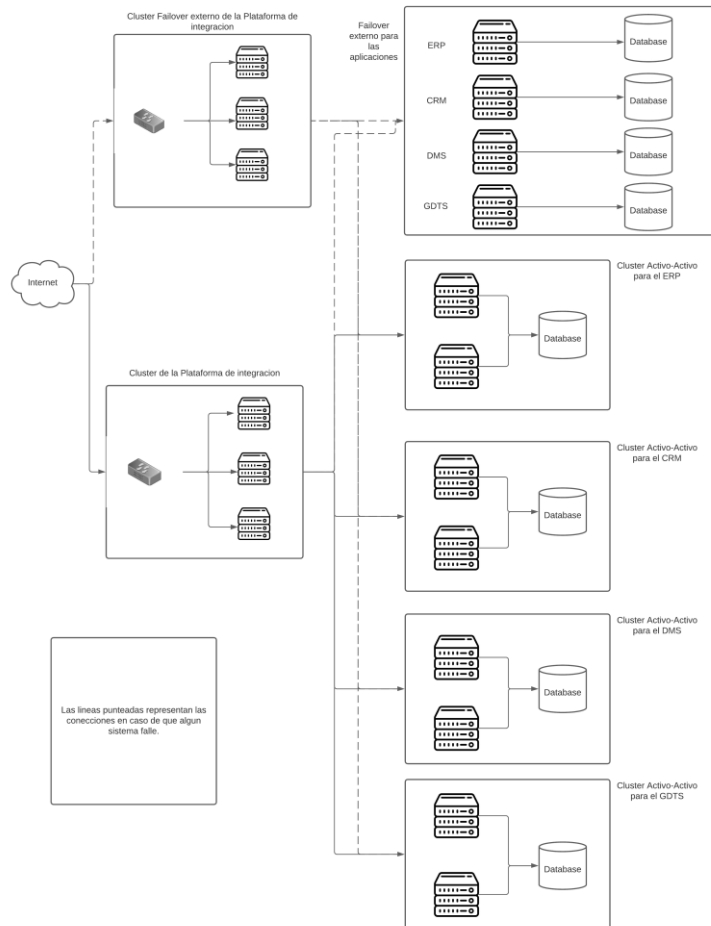
DMS (Escalabilidad): Si aumenta el número de usuarios como se tiene planeado, también aumentara el número de envíos. Por esto es que este sistema tiene que poder escalar, para poder soportar el aumento de envíos que tenga por la llegada de nuevos clientes o por temporadas altas excepcionales donde pocos clientes hacen más pedidos de lo normal.

GDS (Escalabilidad, Desempeño): Con la llegada de las unidades inteligentes esta plataforma va a tener que ser capaz de escalar para poder soportar el incremento de peticiones. Por otro lado, aunque no tiene tantas peticiones como las otras aplicaciones, crear rutas es un proceso pesado y rastrear las unidades puede ser demorado por lo que se necesita que este sistema tenga un muy buen desempeño a pesar de no tener tantas peticiones.

Plataforma de integración (Disponibilidad, Desempeño, Escalabilidad): Este es el punto de entrada para todas las aplicaciones, lo que significa que si falla todo falla. Por esto este tiene que tener una disponibilidad muy alta. Este sistema solo se encarga de pasar peticiones por lo que queremos que a pesar de la gran cantidad de peticiones que le lleguen pueda dirigirlas todas muy rápido, esto para que cuando se mande una petición no se sienta que hay un sistema intermedio entre la app y el cliente.

Finalmente, esta plataforma tiene que escalar junto con todas las aplicaciones del sistema por lo que también es importante que pueda escalar.

2. Infraestructura por aplicación (o grupo de)



(Adjunto se manda esta imagen por

separado)

Proponemos utilizar un cluster por servidores virtuales, ya que permite escalar verticalmente y utilizar un número de múltiples servicios por servidor se vuelve posible. Lo anterior contribuye al desarrollo de la plataforma por integración, lo cual implica que haya una alta disponibilidad. Por esta razón, proponemos utilizar un cluster de failover dada la verticalidad que se quiere. Cabe mencionar que de esta manera conseguimos tener persistencia por almacenamiento compartido en nuestras BDs.

ERP, CRM, DMS, GDTS: Para las aplicaciones, cada una cuenta con un cluster Activo/Activo que comparte una base de datos. Además de esto se tiene un servidor externo para cada aplicación, la base de datos de este será una réplica de la local y esta se actualizará cada 10 minutos para garantizar que no se pierda información.

Plataforma de integración: Para esta se tiene un cluster de varios servidores activos y se tiene un servidor que reparte las peticiones a este cluster. Para garantizar una gran disponibilidad se tiene un cluster idéntico al local, pero de forma remota en caso de un siniestro.

Para las aplicaciones, sabemos que la plataforma de integración cumple con el rol de balanceador de carga

3. Descripción de la infraestructura total

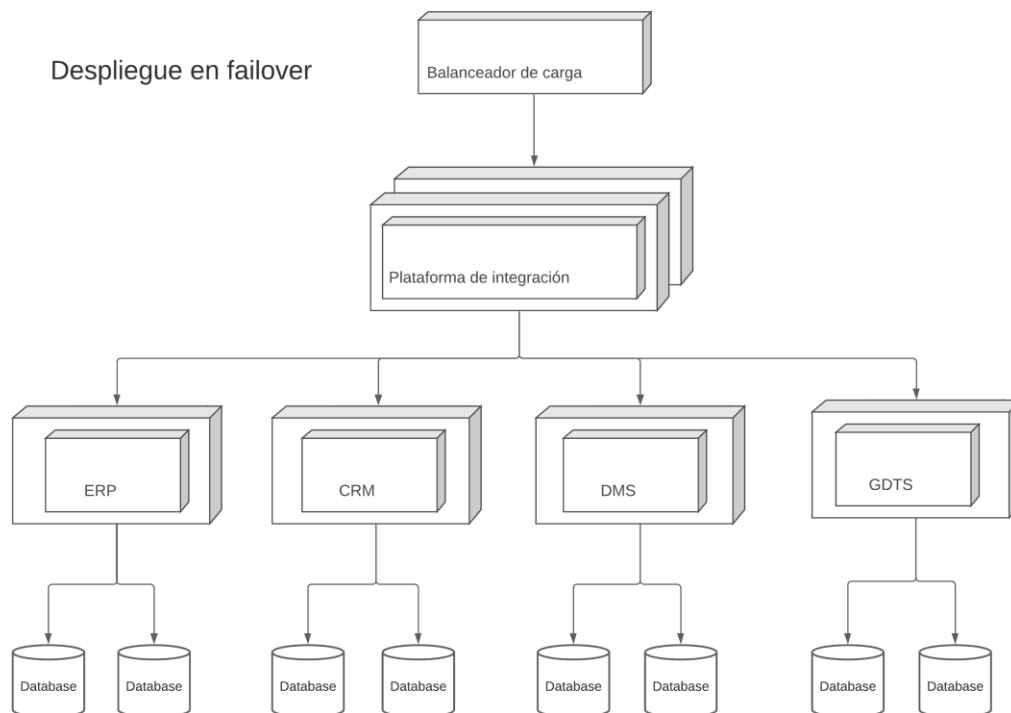
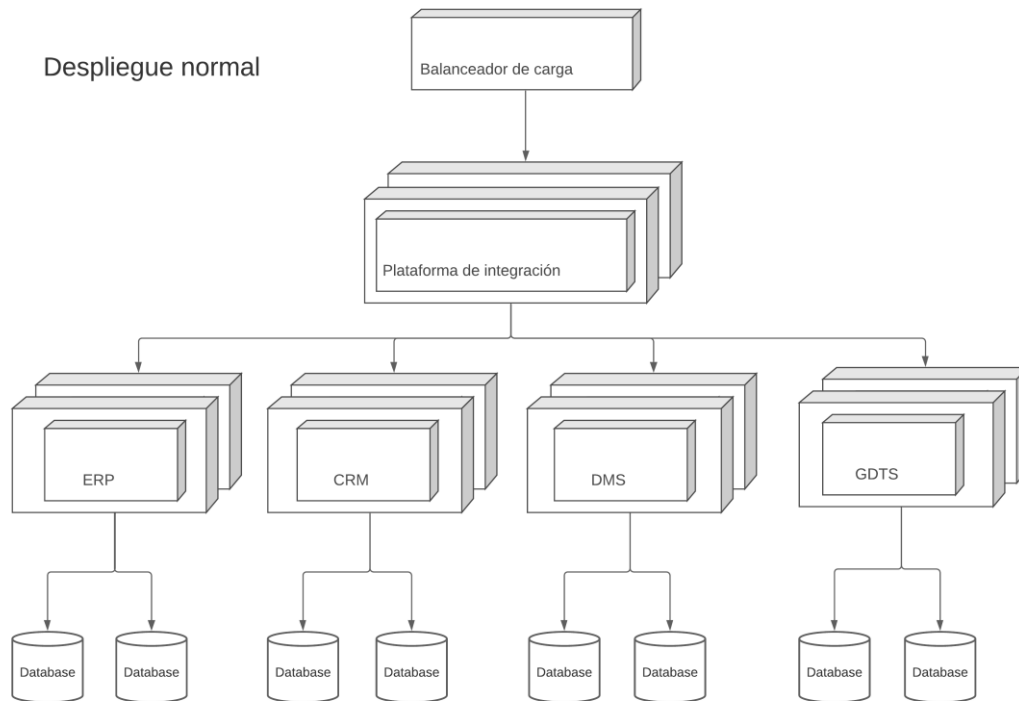
La plataforma de integración se ejecutará en un cluster de tres servidores, en cada servidor al tiempo. Un servidor usará el algoritmo Round-robin para repartir las peticiones a estos tres servidores. Los servidores remotos operarán de forma idéntica y gracias a la virtualización va a poder ser desplegado en poco tiempo en caso de ser necesitado.

La plataforma de integración sabe cada dirección IP de cada servicio por lo que puede distribuir las peticiones de manera acorde.

Todas las aplicaciones están desplegadas sobre un cluster de dos servidores activos que están recibiendo peticiones de la plataforma de integración al tiempo. Estos tienen una base de datos compartida que se encuentra operando en RAID 1 para garantizar tolerancia ante fallos, además su costo de implementación es menor si se compara con el RAID 01 y el RAID 10. En caso de fallo se tiene un servidor externo por cada aplicación. Cada uno de los servidores externos tiene su propia base de datos en RAID 1.

La comunicación entre los componentes del sistema es orquestada por la plataforma de integración que usa un ESB (Enterprise Service Bus) para gestionar esta comunicación. Para manejar la lógica de estos procesos se usa un motor de procesos que administra la consulta de información en el sistema. La lógica es expresada en lenguajes de procesos como BPEL. El sistema CRM se comunica a través de sockets, mientras el sistema ERP se comunica a través de servicios web basados en SOA y a través de la administración del mecanismo comunicación. El sistema DMS, por su parte, se comunica a través de la gestión de archivos.

En general todas las peticiones llegan a un balanceador de carga a través de https, este tiene un firewall y la lógica para distribuir las peticiones a cada réplica de la Plataforma de Integración de manera apropiada. Una vez que la plataforma tiene la petición, la lee para saber que recurso es el indicado para responder a dicha petición. Luego procede a pasar la petición por http a la aplicación indicada la cual la recibe y la resuelve. Si es necesario, la aplicación se comunica con la base de datos y/o con otras aplicaciones por el protocolo que corresponda.



Esta configuración puede verse un poco grande para el GDTS, pero como esta es el que se espera que incremente más en peticiones en los siguientes años, eventualmente se necesitara esta plataforma.

En síntesis, la infraestructura que queremos debe interconectarse a través de un ESB, que es el encargado de envolver los protocolos de comunicaciones entre aplicaciones en servicios estandarizados con operaciones claramente definidas, cada un operando con tipos de datos canónicos respectivos a cada sistema de información.