

Arquitecturas de Alta Disponibilidad y Escalabilidad

Ferran Garcia Pagans
Ferran.Garcia@Oracle.com
Principal Sales Consultant



Agenda

- Introducción
 - Grid Computing y HA.
- Topología de Sistemas, Escalabilidad y HA.
 - Componentes que forman parte de un sistema informático y problemas de HA.
- Pérdidas de Servicio
 - Tipos de Pérdidas de Servicio y soluciones.
- Mundo Senior, un caso real.
- Conclusiones



Introducción



Arquitectura Grid Computing

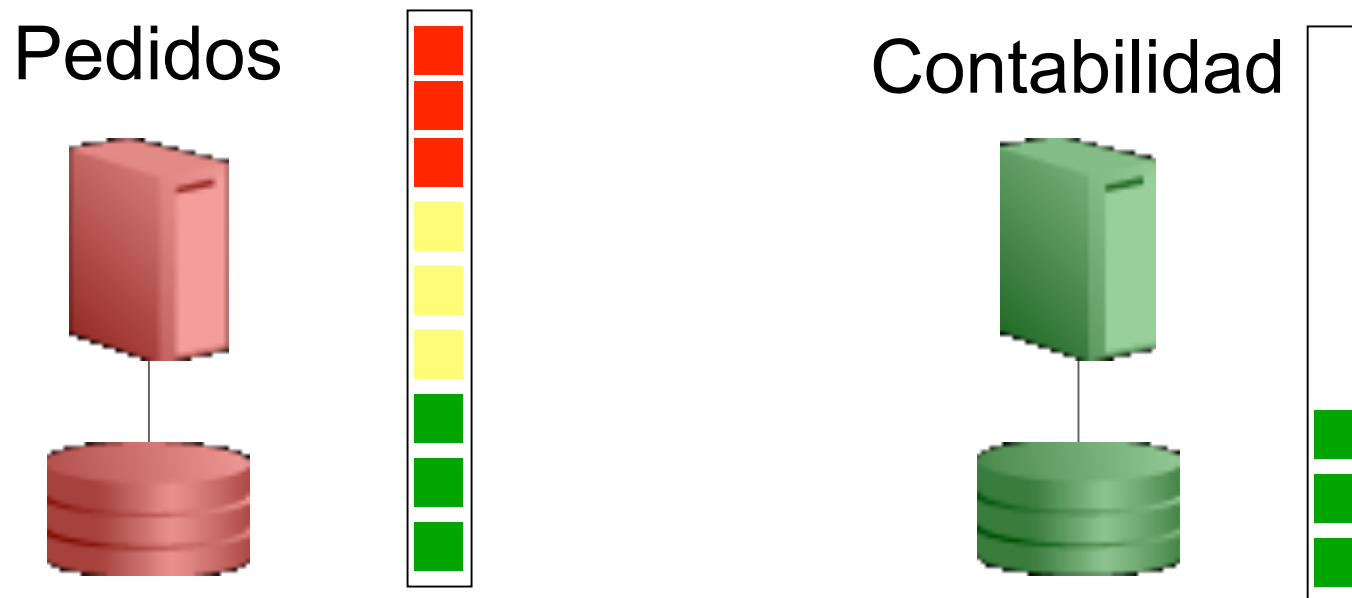
Grid empresarial

- Uso coordinado de muchos servidores pequeños actuando como un gran ordenador.
- Permite reducir el TCO y mejorar la QoS.
 - Estandarización de la plataforma HW
 - Plataformas HW de bajo coste
 - Escalabilidad y Alta Disponibilidad a un coste reducido
 - Estandarización del mantenimiento



Ejemplo: Diciembre

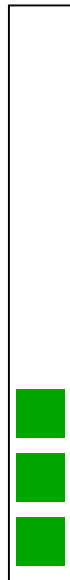
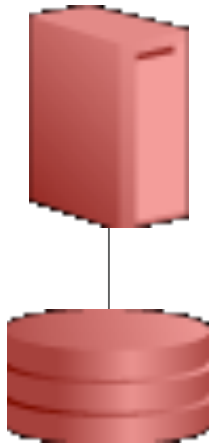
- Gestión de pedidos llega a su pico de carga
- Aplicación financiera en bajo uso de capacidad



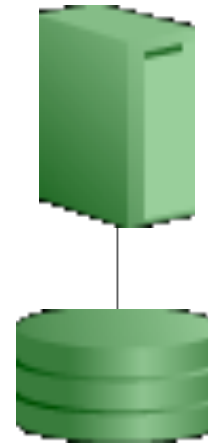
Ejemplo: Enero

- Gestión de pedidos pasa a carga baja
- Aplicación financiera en pico de carga por cierre del año

Pedidos



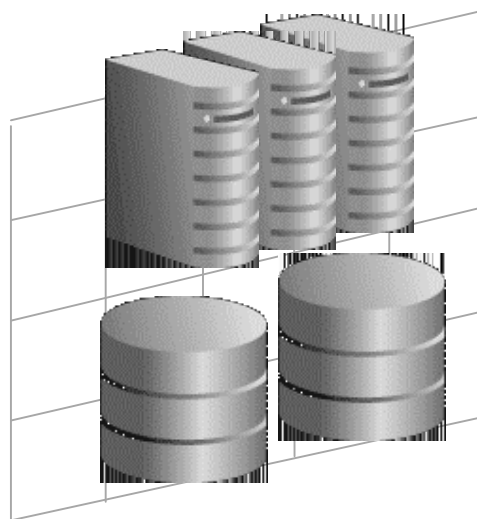
Contabilidad



Ejemplo: Consolidando ...

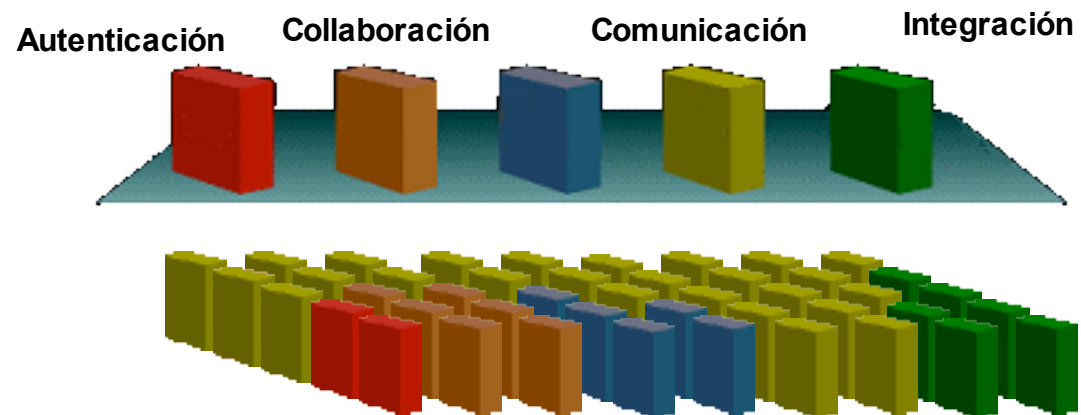
- Balance de carga para optimizar los recursos en ambos picos de carga

Pedidos y Contabilidad



Escalar con HW de Bajo Coste

- **Uso de Recursos Eficiente**
- **Añadir Recursos Cuando Sea Necesario**
- **Escalar Servicios y Apps de Forma Independiente**
- **Todas las Plataformas**
- **Todas las Arquitecturas**



Virtualizar y Pool de Recursos

La HA es lo que permite que el Grid funcione!



Beneficios

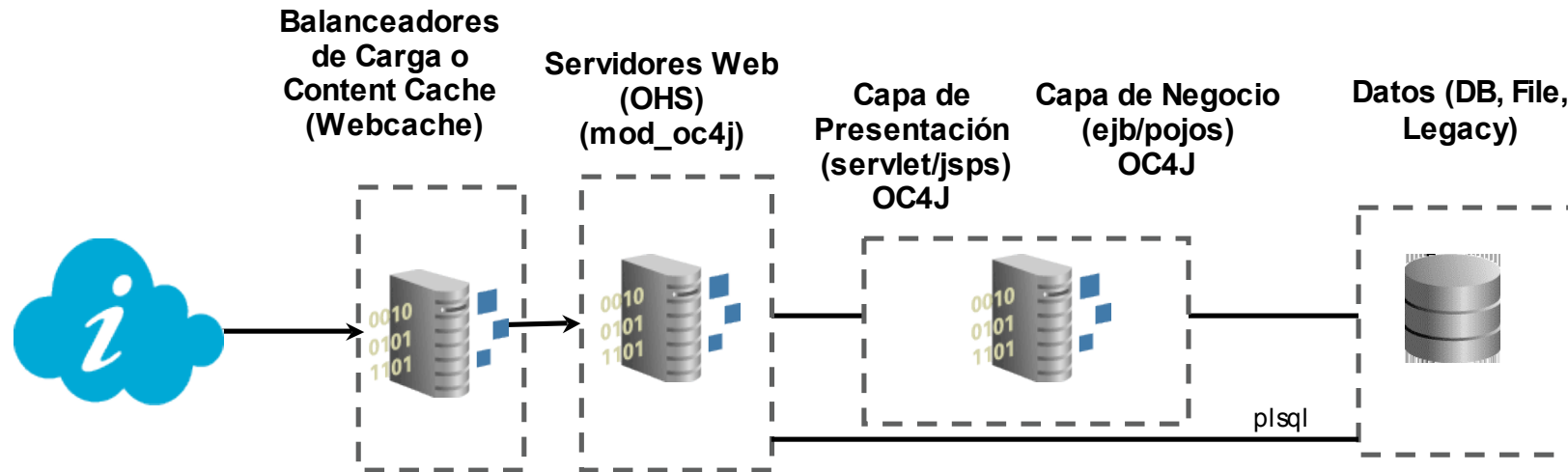
- Mejora de la calidad de servicio.
 - Rendimiento.
 - Escalabilidad, capacidad bajo demanda.
 - Alta disponibilidad.
- Reducción del coste de propiedad (TCO).
 - Hardware de bajo coste.
 - Reducción necesidades de administración.



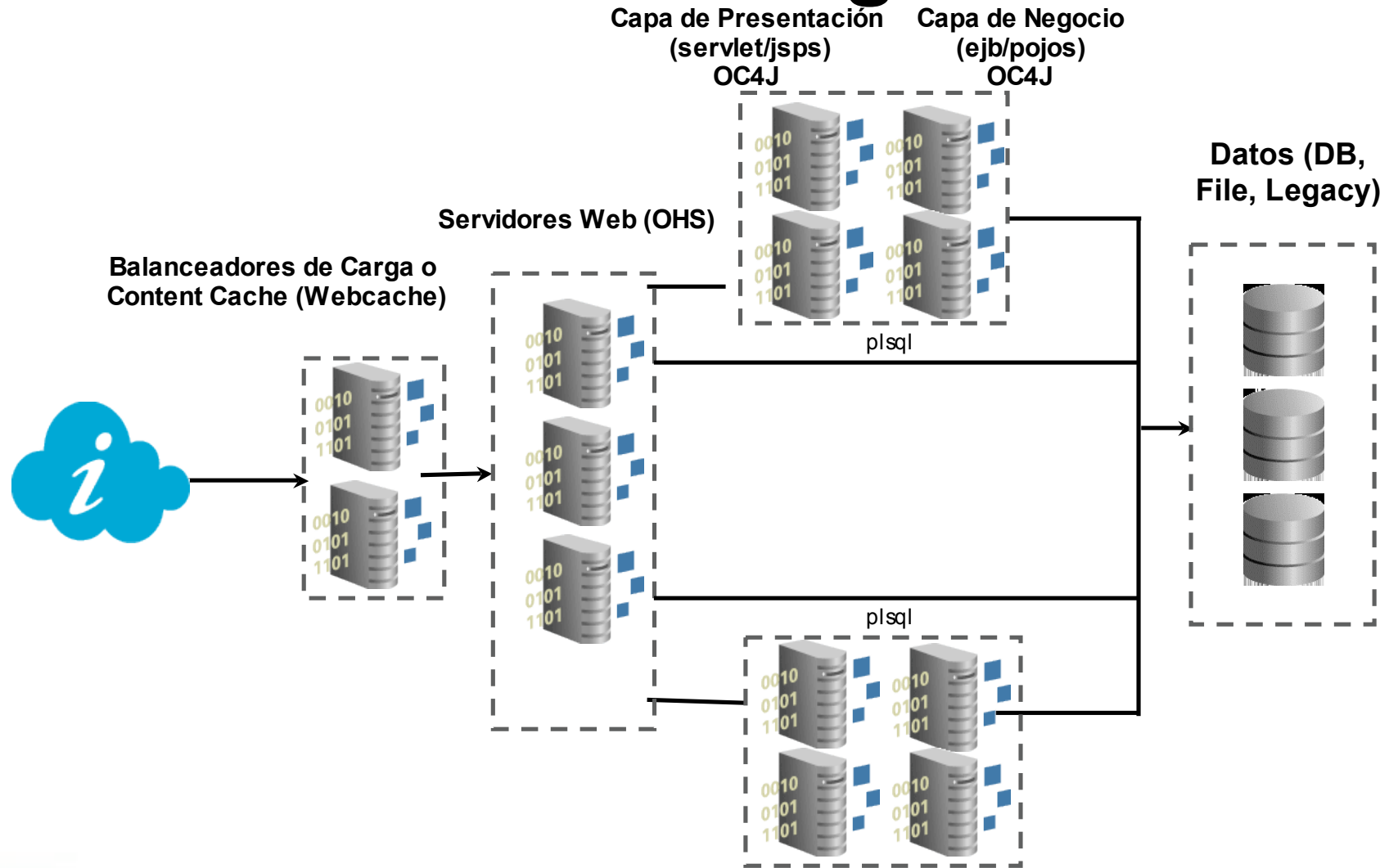
Topología de Sistemas, Escalabilidad y HA



Oracle Application Server



Clustering

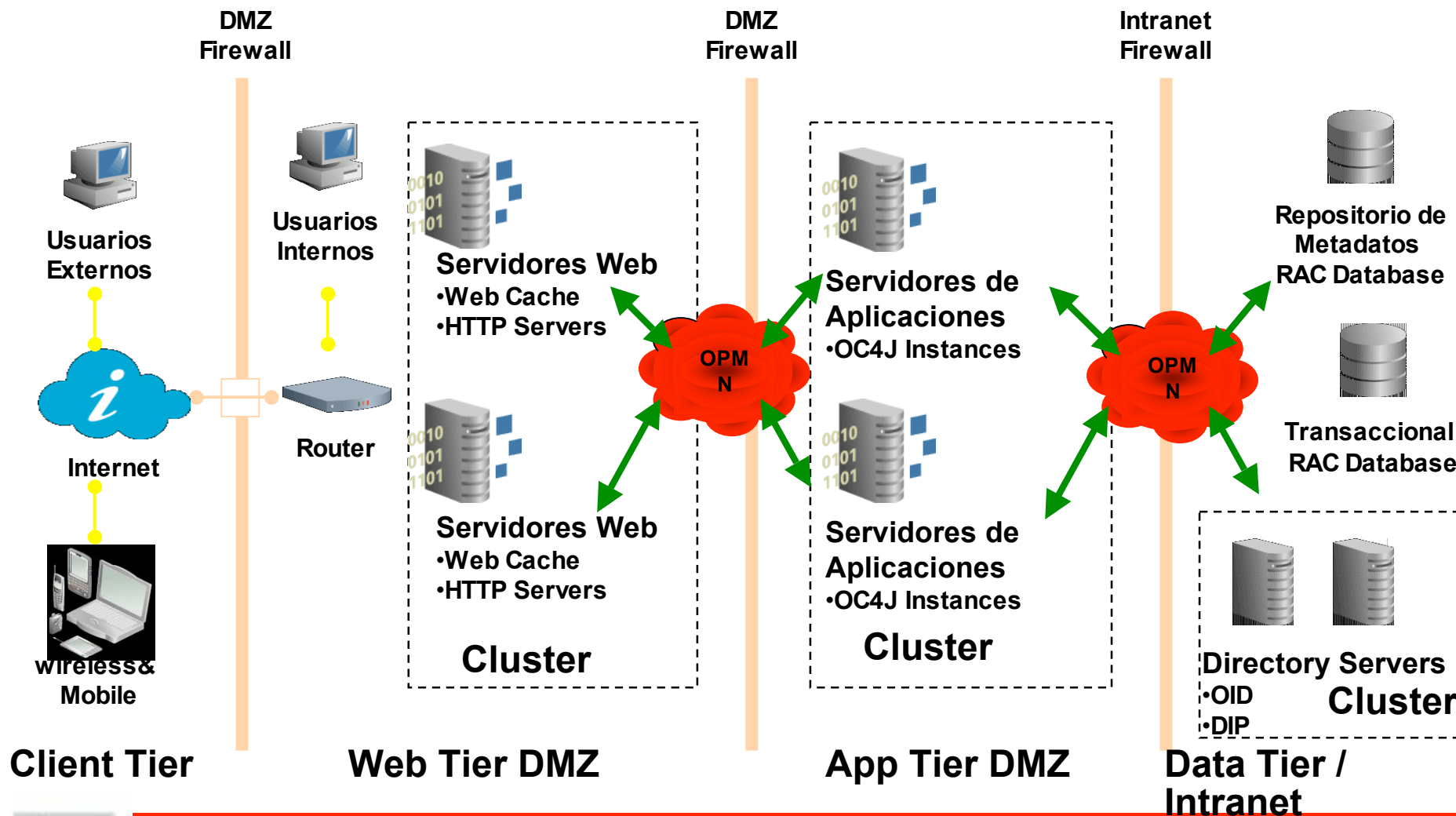


Puntos críticos

- Optimización del Código.
- Balanceo de Carga.
 - Detección de caídas.
 - Re-arranque de componentes.
 - Protocolos de balanceo de carga.
 - Tiene un impacto importante en el rendimiento de la red, en la capacidad de recuperarse de fallos y en la de evitarlos..
- Replicación de sesión.
 - Afecta al rendimiento del HW y de la red y a la capacidad de mantener sesión.
- Evitar un punto de fallo único.
 - El sistema es tan robusto como lo es el punto menos robusto del sistema.



Topología de un Sistema Web



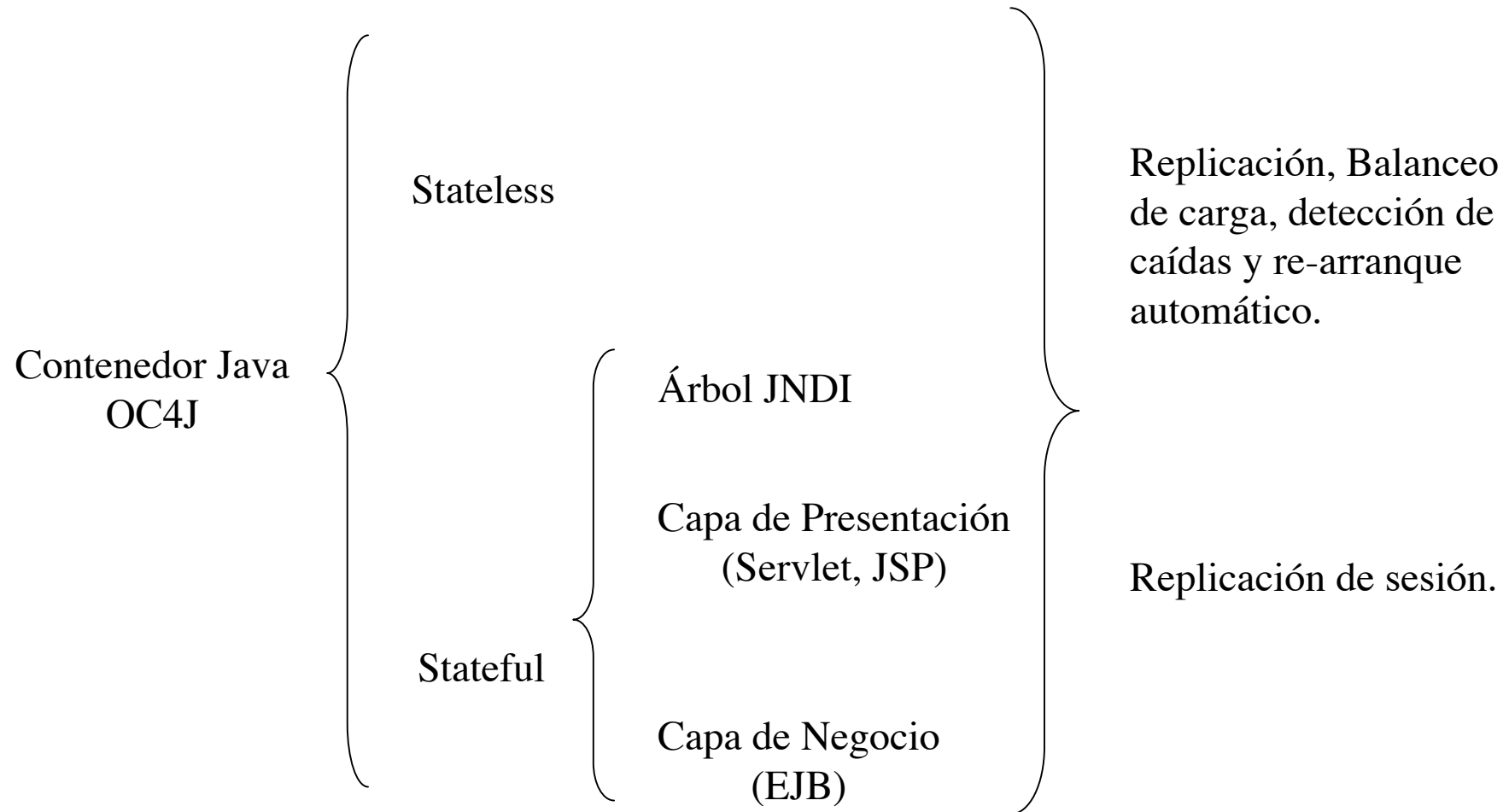
Balanceo de Carga

Componentes	Criticidad (0-5)	Mecanismos de Balanceo
Webcache y Servidores Http	3	Round-robin, weighed round robin
Servidores Http y Contenedores J2EE	5	Random, Round Robin, Random with Local Affinity, Round Robin with Local Affinity, Random using Routing Weight, Round Robin using Routing Weight, Metrics Based, Metric Based with Local Affinity
Intra Contenedores J2EE	3	Random
Contenedores J2EE y base de datos	5	Mecanismos de Oracle Net (RAC to the rescue)

datos



Contenedor Java/J2EE - HA



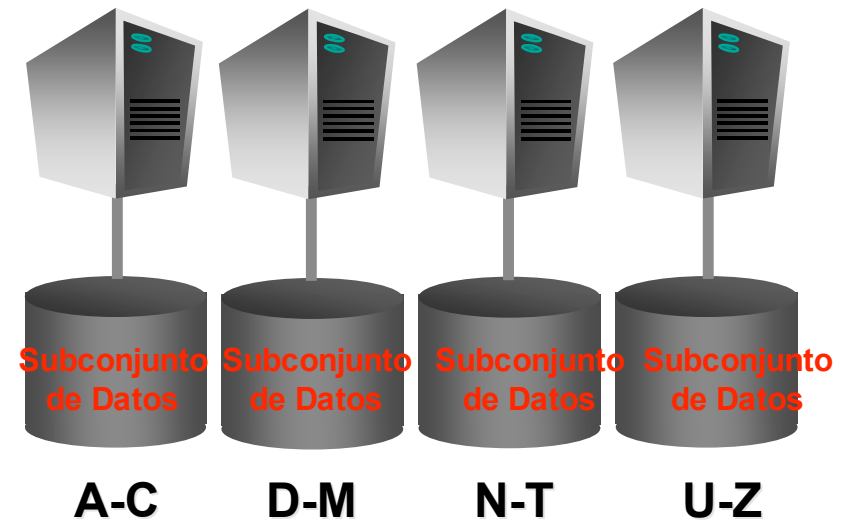
Políticas de Replicación

- Definen cuando se replican los atributos y las variables.
- Cuando la replicación es demasiado frecuente produce pérdidas de rendimiento en la red y los servidores.
- Si la replicación es muy poco frecuente puede producir pérdidas de datos.
- Posibilidad de persistir la sesión en base de datos.

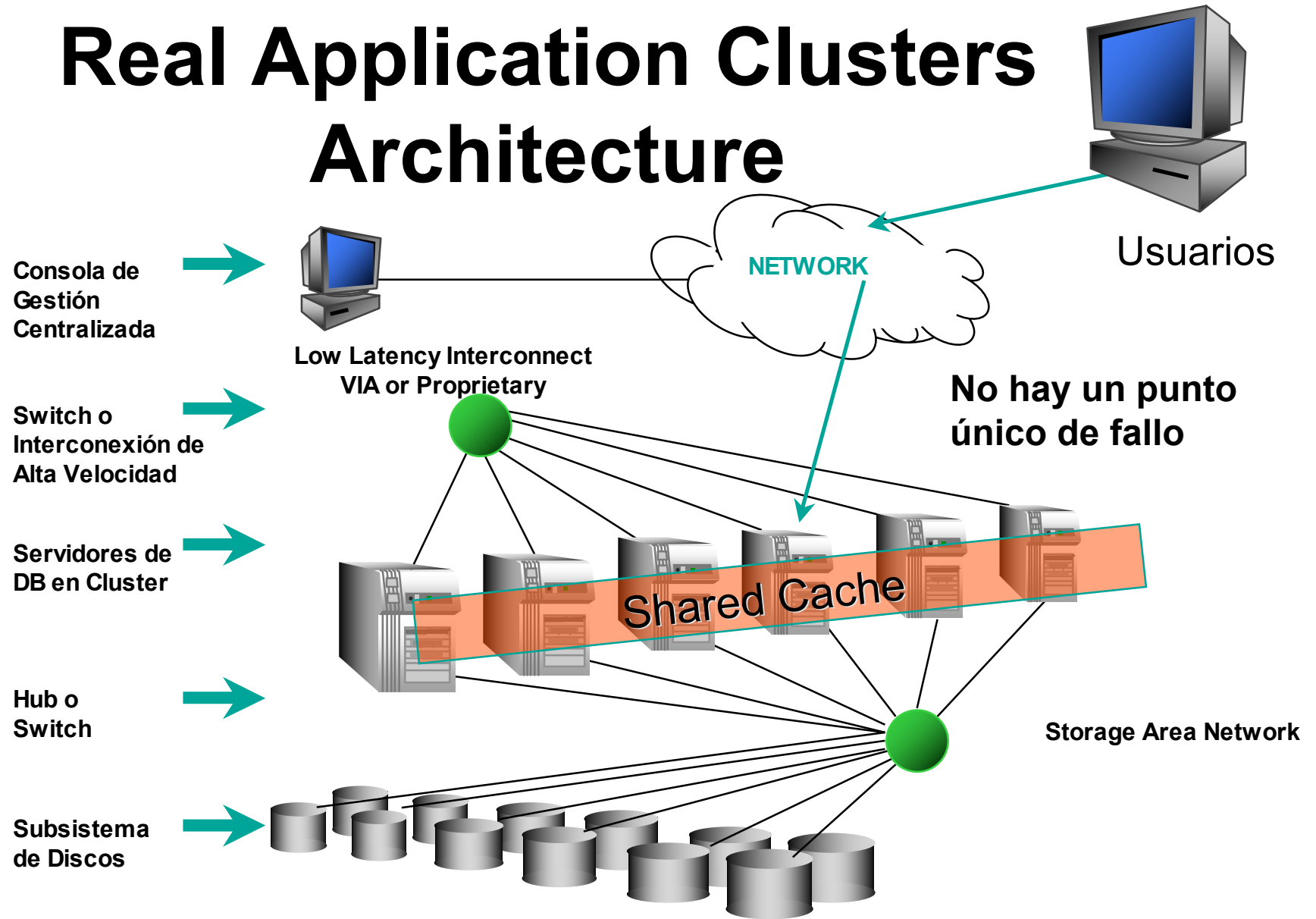


Bases de Datos en Cluster

- ‘Shared Disk’
 - Real Applications Cluster
 - DB2 para Mainframes
- ‘Shared Nothing’ y Federadas

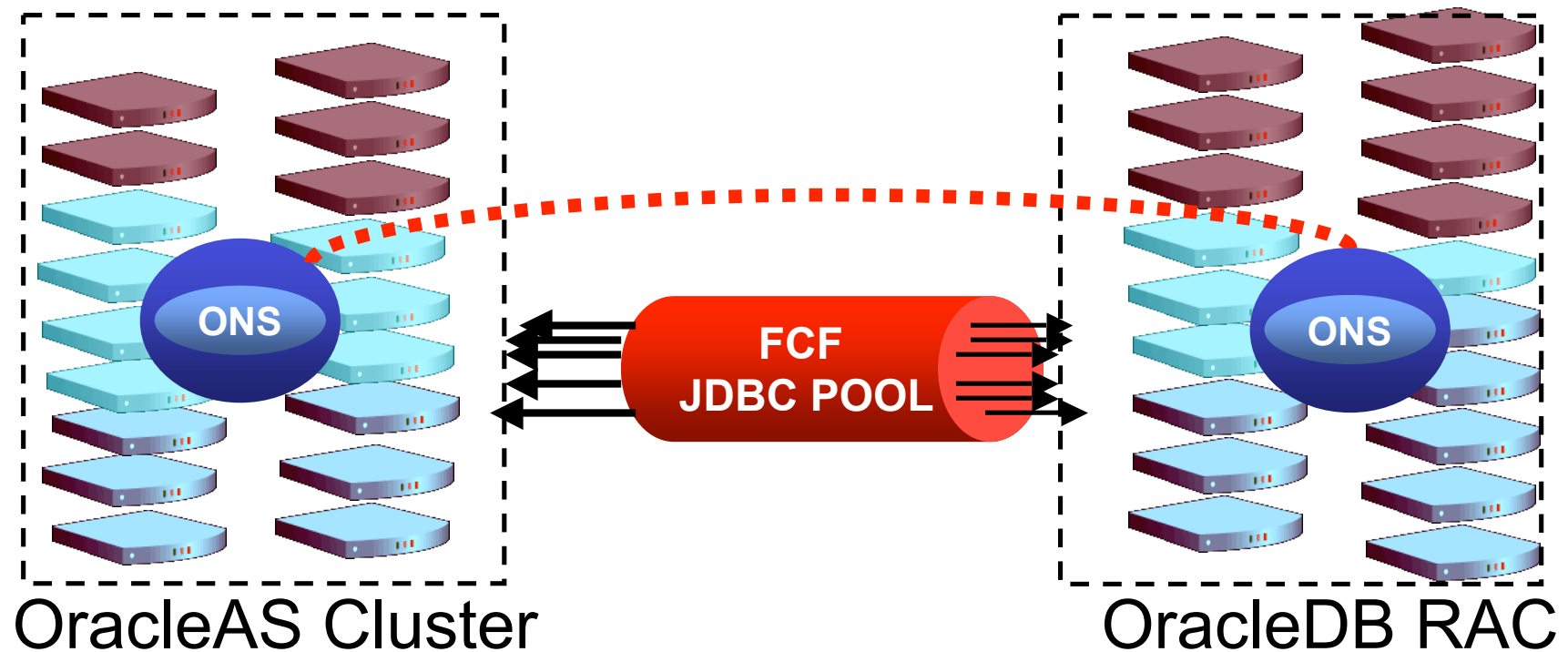


Real Application Clusters Architecture



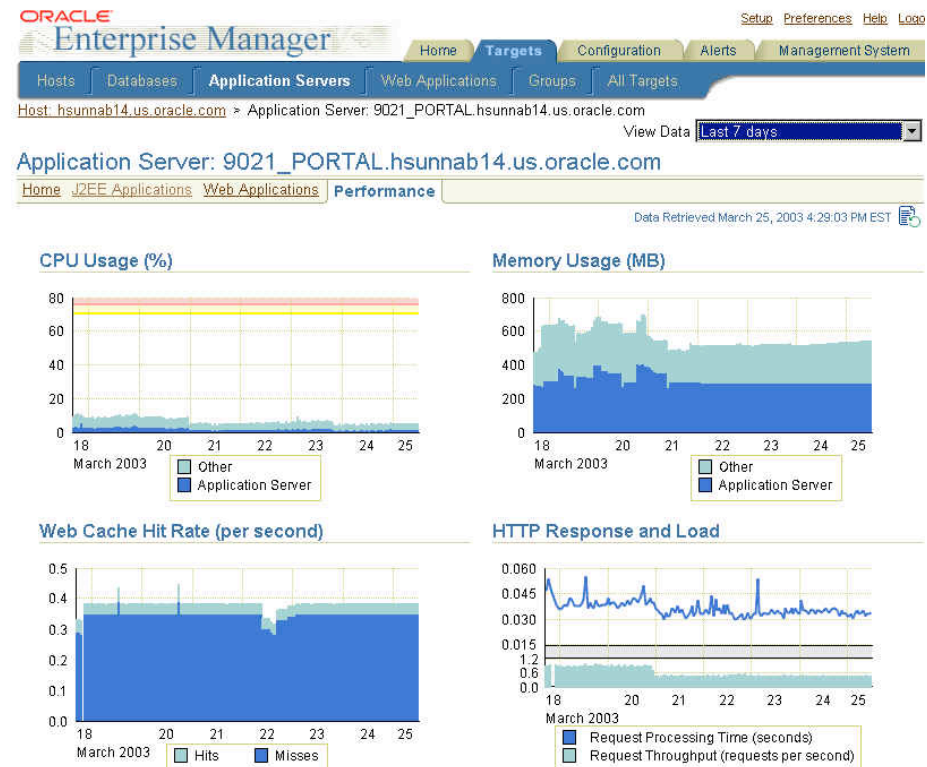
Data source- HA

- Enlazar las bases de datos en Grid con el servidor de aplicaciones en Grid



Administración y Monitorización con Enterprise Manager

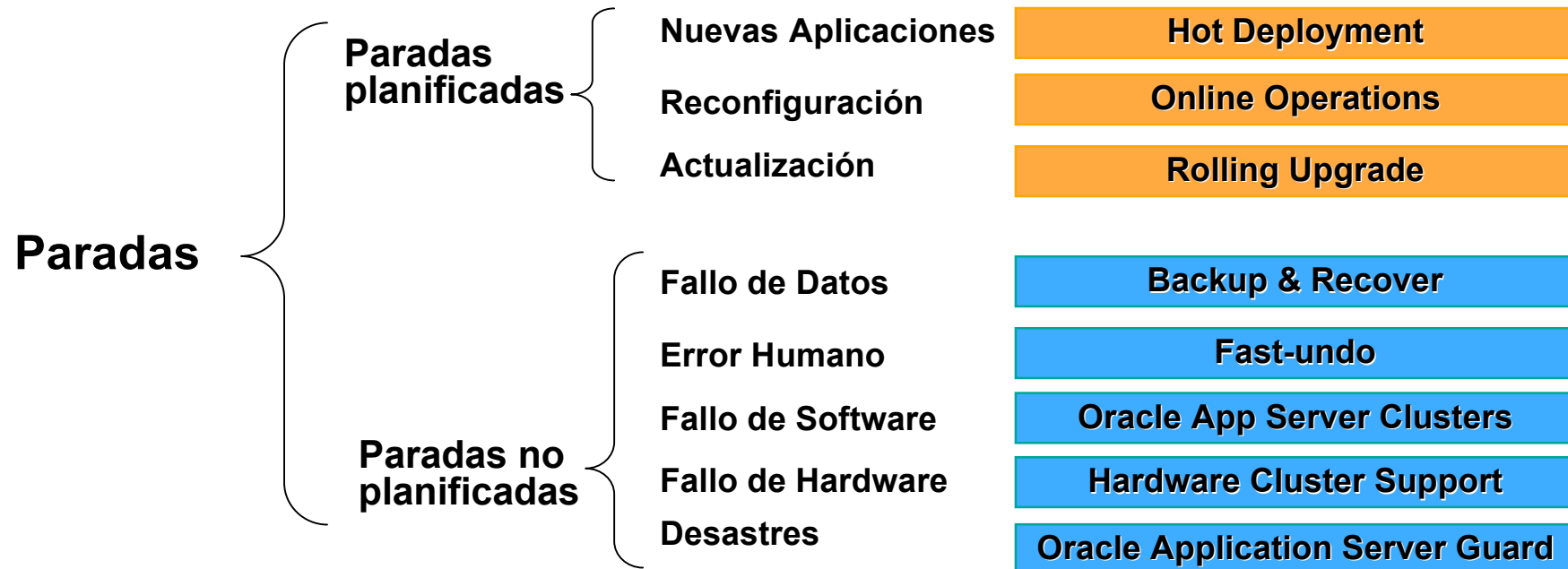
- Admin. y Monitorización de todo el entorno Oracle
- Recoge métricas de: CPU, Memoria, HTTP, componentes Oracle, ...
- Gestión de parches
- Clonación de instancias
- Gestión de alertas
- Control de Transacciones
- Etc...



Pérdidas de Servicio



Paradas de los Sistemas



Tipos de Clusters

	Oracle Application Server Clusters	Oracle Application Server Cold Failover Clusters
Escalabilidad	N servidores de aplicación trabajando en paralelo.	1 servidor de aplicaciones atiende las peticiones.
Coste de Operación	Gestionar un Cluster.	Gestionar un servidor de aplicaciones.
Coste de Licencias	Licencia por CPU's	Licencia por CPU's activas.
Alta Disponibilidad	Generalmente es más transparente al usuario.	
Instalación	Es necesario instalar N servidores de aplicaciones.	Se instala 1 servidor de aplicaciones y se clona.



Diseño de Alta Disponibilidad

Requisitos de Negocio			Selección Arquitectura	
HA Local	Escalabilidad	Protección de Desastres	Instancias Redundantes	Recuperación de Desastres
x	x	x	Básico	x
✓	x	x	Activo - Pasivo	x
x	✓	x	Activo – Activo	x
x	x	✓	Básico	✓
✓	✓	x	Activo – Activo	x
✓	x	✓	Activo - Pasivo	✓
x	✓	✓	Activo – Activo (MT) Básico (Infra)	✓
✓	✓	✓	Activo – Activo (MT) Activo - Activo (Infra)	✓





Un Ejemplo Real

MundoSenior, MundoRed



MundoRed – Mundo Senior

Empresa participada por:

- Viajes Marsans
 - Viajes Iberia
 - Viajes Barceló
 - Halcón Viajes
-
- Concesión del programa de viajes del IMSERSO.



Factores de Cambio

Se decide reemplazar el sistema informático no propio y adaptado tras 12 años.

- Acceso a todas las agencias de viajes.
- Tener el inventario en sus propios servidores.
- Aplicación desarrollada para su negocio.
 - Particularidades funcionales.
 - Picos de carga.
- Actualización tecnológica.



Escuchar al cliente

“necesitábamos un sistema muy estable, robusto y redundante, ya que nuestros servicios han de ser capaces de arrancar a las 9:00 horas de la mañana, desde que se ponen a la venta las plazas. Se debe pensar que a partir de esa hora, en los primeros 15 segundos todas las agencias de viajes necesitan realizar una reserva. A partir de ahí, y en la primera media hora, el sistema debe funcionar a un ritmo de diez reservas cerradas y con emisión de documentación por segundo.”



Factores Críticos de Éxito

- Rendimiento.
 - 10 Reservas cerradas y con emisión de documentación por segundo.
 - Código, Afinamiento de la plataforma.
- Alta Disponibilidad.
 - Evitar tener punto únicos de fallo.
 - Redundar todos los componentes de la plataforma.
- Escalabilidad.
 - Crecimiento del negocio, nuevas líneas de productos.
- Flexibilidad.
 - El entorno presenta picos de carga muy pronunciados.
 - Nuevas líneas de negocio.



Diseño de Alta Disponibilidad

Requisitos de Negocio			Selección Arquitectura	
HA Local	Escalabilidad	Protección de Desastres	Instancias Redundantes	Recuperación de Desastres
x	x	x	Básico	x
✓	x	x	Activo - Pasivo	x
x	✓	x	Activo – Activo	x
x	x	✓	Básico	✓
✓	✓	x	Activo – Activo	x
✓	x	✓	Activo - Pasivo	✓
x	✓	✓	Activo – Activo (MT) Básico (Infra)	✓
✓	✓	✓	Activo – Activo (MT) Activo - Activo (Infra)	✓



Participación de Oracle

- Proyecto liderado por Brújula
- Funciones Consultivas
 - Definición de la Arquitectura
 - Monitorización y optimización del aplicativo
- Servicios Profesionales – APS/ACS
 - Pruebas de stress
 - Verificación Instalación
 - Soporte on-site a la puesta en marcha



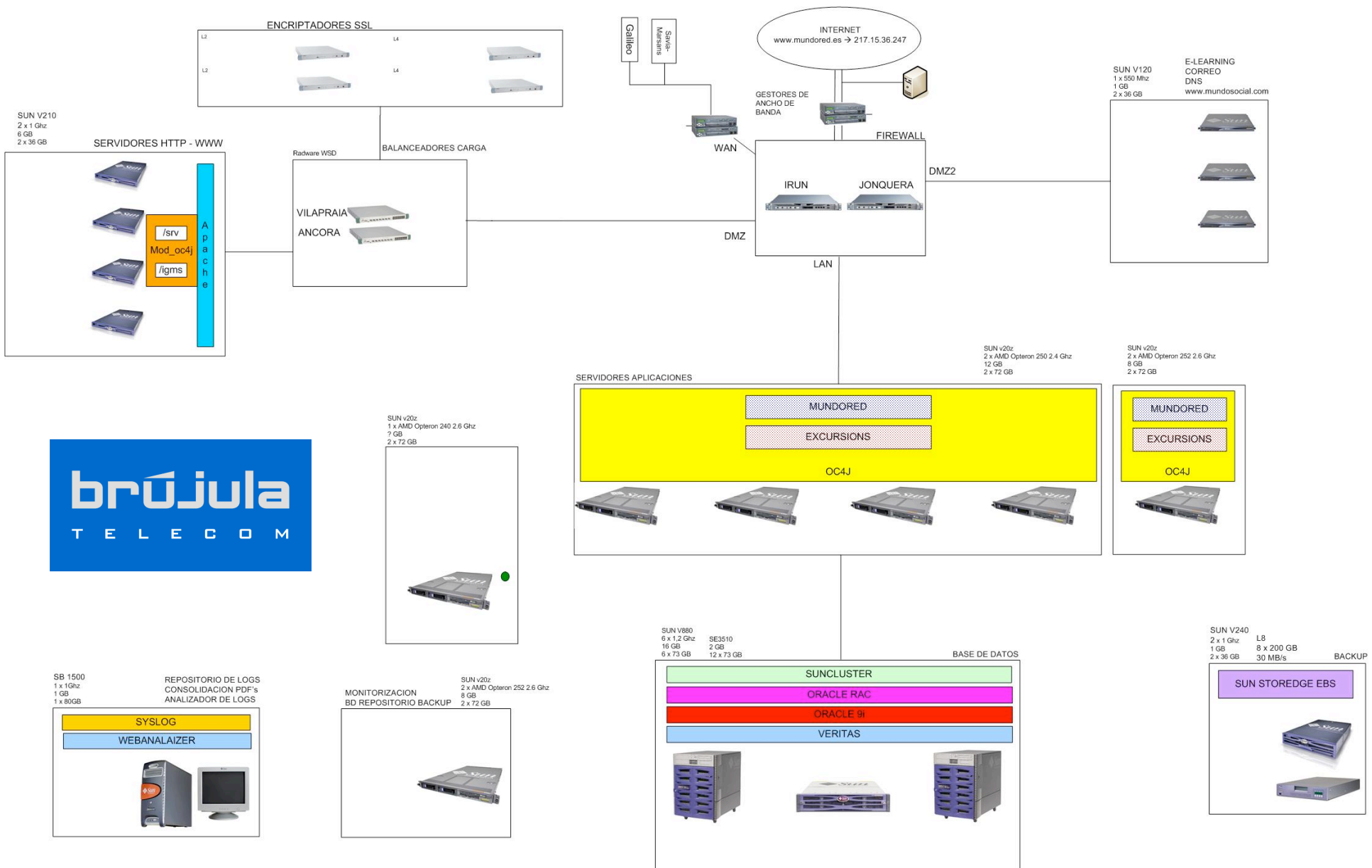
Problemas de Rendimiento

```
public class CreaReservaSessionEJB implements SessionBean{  
public Object[] reserva(....)
```

1. validateRestriccionesPasajeros
2. actualizaAllCupoSalidaVOs
3. actualizaAllCupoSalidaVOs

- Debería ejecutarse 10 veces por segundo!!!
- En los primeros tests algunas veces llegaba a tardar 3 segundos en ejecutarse.





Resultados

- 13 Reservas por segundo!!!!
- 22,000 reservas (51,000 plazas) en media hora!!!
- Media de sesiones conectadas al sistemas 6,500
- Incremento de 1,000 puntos de venta.
- 100% Internet, sin instalaciones.
- Sistema más ágil y flexible.



Conclusiones



Rendimiento y HA

- Muchos problemas de rendimiento se deben a la programación o diseño de los datos (Java, SQL ,...).
- La regla de Oro es evitar un punto de fallo único.
- Analizar las políticas de replicación de sesión y balanceo de carga.
- La HA tiene un coste económico directo.
- Las pruebas son fundamentales, debería hacerlas un equipo independiente.
- El equipo humano es fundamental!!!!
 - ¿Quién me va a dar soporte a esta tecnología?



Arquitecturas de Alta Disponibilidad y Escalabilidad

Ferran Garcia Pagans
Ferran.Garcia@Oracle.com
Principal Sales Consultant

