



JP ELEKTROPRIVREDA BiH

d.d. - Sarajevo

24.hrOUG
godišnja konferencija

IZAZOVI ODRŽAVANJA I UNAPRIJEĐENJA BAZA PODATAKA I VIRTUALNIH MAŠINA NA ODA-ma

Belma Štukan, dipl.ing.el.

Elvir Hadžić, dipl.mat.i informatičar









Sadržaj

- Razlozi za migraciju na ODA-u
- Organizacija EPBiH
- Organizacija i karakteristike ODA-e
- Hardverska arhitektura
- Osnovne prednosti ODA-e
- Automatic Storage Management
- Oracle ASM instance/grupe
- Oracle ACFS
- ODA virtualizirana platforma
- Oracle ZFS Storage appliance
- Q&A

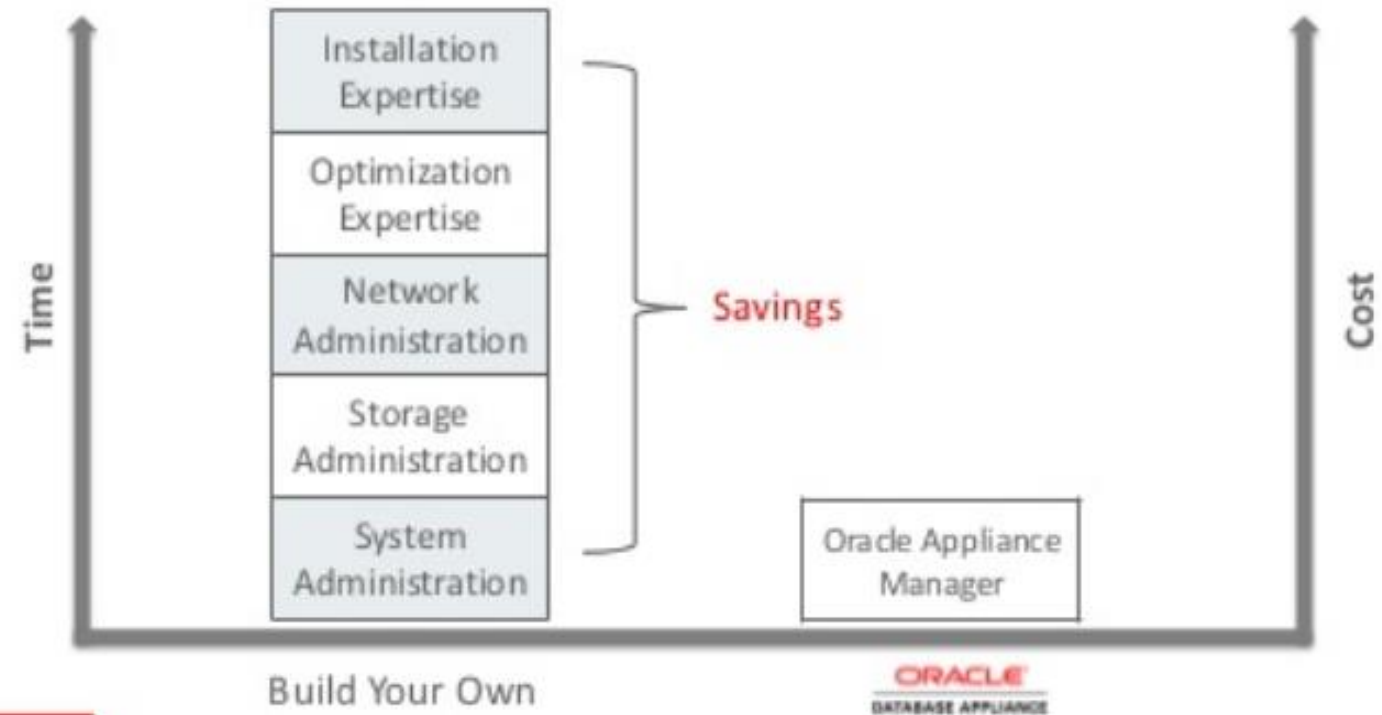
Razlozi za migraciju na ODA-u

- Jako veliki broj baza podataka na različitim sistemima – različit hardver i različiti operativni sistemi
- Jedna hardverska platforma za sve baze podataka
- Jednostavnija instalacija i podešavanja
- Jednostavnije održavanja cijelog sistema (administracija, patch-iranje, storage,...)
- Pojednostavljen razvoj i podrška za visoko dostupna database rješenja
- Raspodjela licenci

Compare to Build Your Own

Build Your Own		Oracle Database Appliance	
5 Puzzle Pieces • Server, storage, networking, database, consultants		1 Component • Push one button for install	
7 Staff / Skills for HA • DBA, network admin, storage admin, system admin, installation expertise, HA expertise, optimization skills		1 DBA	
24 + Patches Per Year		4 Patches Per Year	
~2,200 Hours Maintenance		~ 110 Hours Maintenance	

To Install, Manage and Maintain



PROIZVODNJA, SNABDIJEVANJE I DISTRIBUCIJA

Distributivni centri u EPBIH

- PTZ Sarajevo
- ED Zenica
- ED Tuzla
- ED Mostar
- ED Bihać



Organizacija i karakteristike ODA-e

- Oracle database appliance (ODA) je organizovana sa ugrađena 2 servera sa zajedničkim storage-om
- Omogućuje pojednostavljenje razvoja, održavanja i podrške za visoko dostupna database rješenja
- Oslanja se na Oracle Real Appliaction Cluster (Oracle RAC)
- Odvojeni prostori i načini da je moguće kreirati virtuelne mašine na svakom od ova dva servera i/ili da se iste kreiraju na zajedničkom storage-u
- Ukupan memorijski prostor je moguće podijeliti na dijelove za baze i na dio za virtualne mašine



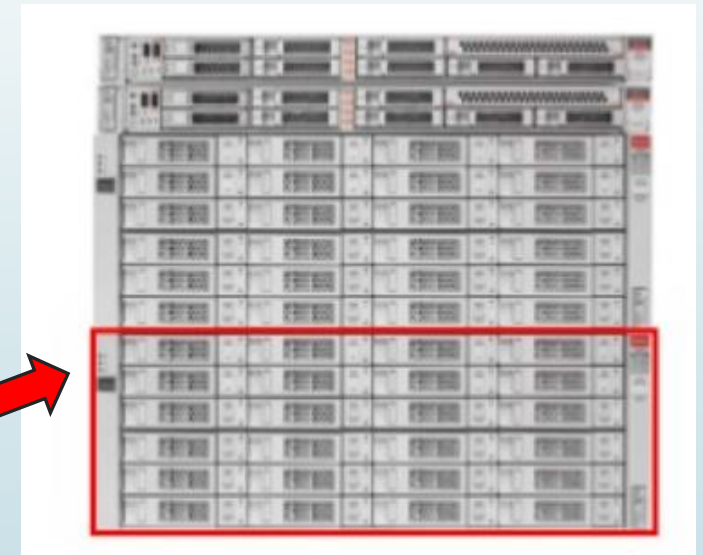
Organizacija i karakteristike ODA-e

- Dva serverska čvora interno su povezana redundantnom 1GbE vezom za komunikaciju klastera
- Uz pomoć Appliance Managera moguće je definirati broj jezgri procesora koje se aktiviraju kako bi se direktno utjecalo na troškove licenciranja softvera
- Verzije sa opcijom HA (high availability) omogućuju podršku za Big Data u smislu podrške sa prostorom te sa opcijama za podršku za Enterprise Oracle baze podataka



Hardverska arhitektura (ODA)

- Dva servera, svaki sa:
 - 36 Intel Xeon CPU core-ova
 - 256 GB memorije, moguće proširiti do 768 GB
 - Redudant InfiniBand Interconnect
 - 10 GbE Public Network
- Storage Shelf:
 - 800 GB flash log storage
 - 1.6 TB flash cache storage
 - 64 TB HDD
- ODA X5-2 dodatni Expansion Shelf
 - Dodatnih 64 TB HDD, 128 TB za DATA
 - Dodatnih 800 GB SSD, 1.6 TB za REDO
 - Dodatnih 1.6 TB SSD, 3.2 TB za FLASH



Osnovne prednosti ODA-e

- Jednostavnost u instalaciji i konfiguraciji
- Lakše upravljanje rješenjem zbog integrisanog patch-iranja
- Podrška za rješavanje problema
- Skalabilnost rješenja u smislu da je moguće aplicirati/koristiti od 2 do 72 procesorska core-a odnosno licencirati onoliko koliko koristimo resurse
- Postojanje odvojenih grupa diskova za različite tipove database objekata

Nedostaci evidentirani u toku korištenja ODA platforme:

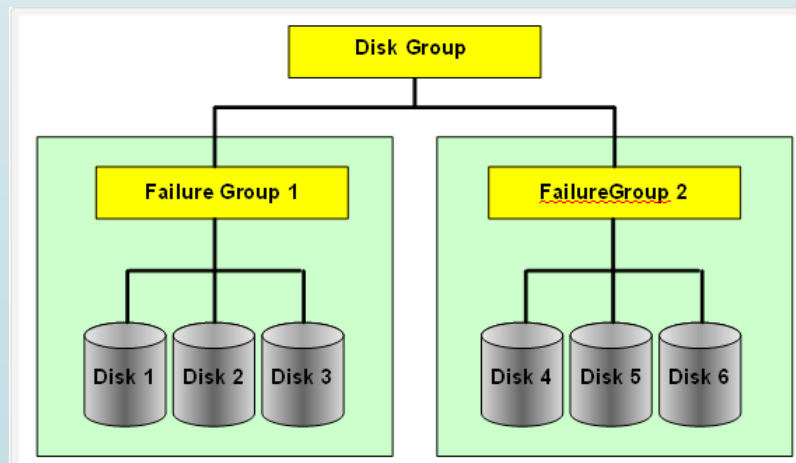
- Nemogućnost korištenja nekih opcija za ASM raspoloživih kod ručno pripremljenih rješenja

- Prije instalacije i konfiguracije ODA platforme potrebno napraviti plan instalacije:
 - hardverska bare-metal opcija načina instalacije (konfiguracija ODA-u bez podrške za virtualne mašine)
 - virtualna opcija načina instalacije (konfiguracija ODA-u sa podrškom za virtualne mašine)
- U slučaju odabrane opcije bare-metal kompletna ODA platforma će biti dedikirana za rad isključivo Oracle baza podataka
- U EPBIH postoje ODA-e sa oba načina konfiguracije

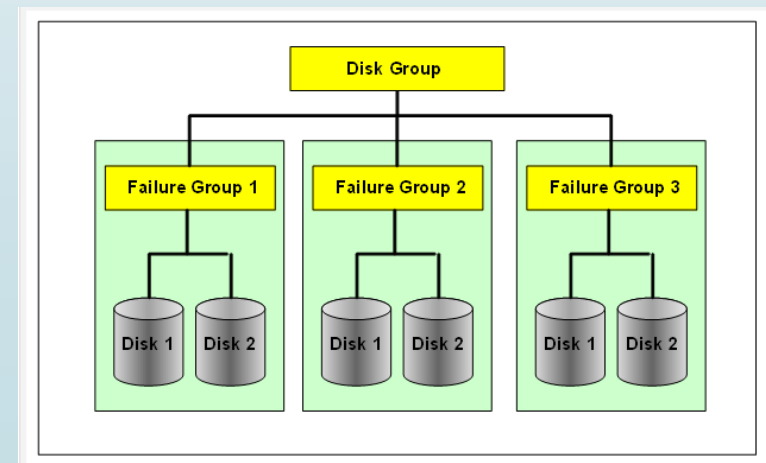
Automatic Storage Management (ASM)

- Oracle-ovo preporučeno rješenje za upravljanje prostorom
- Koristi disk grupe za pohranjivanje podataka
- Oracle ASM normalna i visoka redundancija disk grupa omogućava dvostruki ili trostruki mirror-ing

Normal – 2-way mirroring



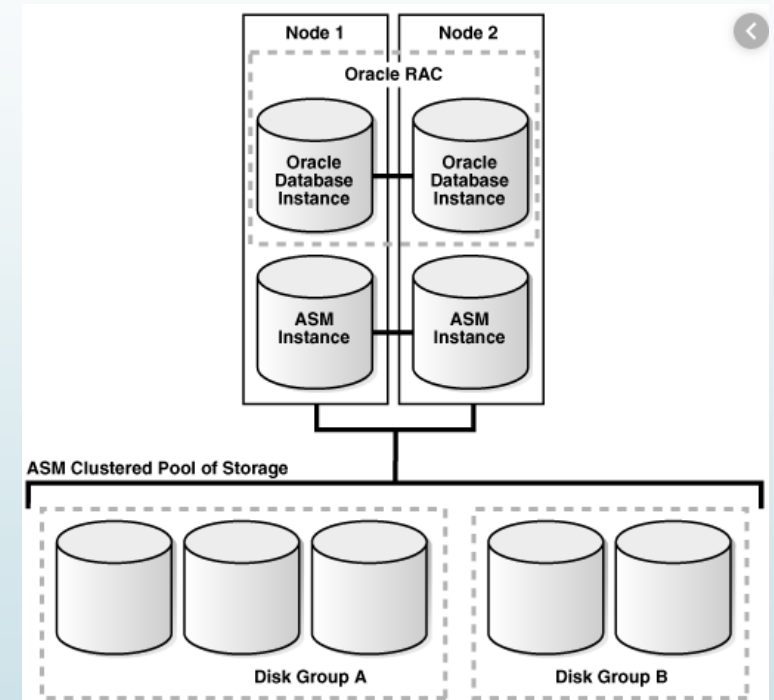
High – 3-way mirroring



Oracle ASM instance/grupe

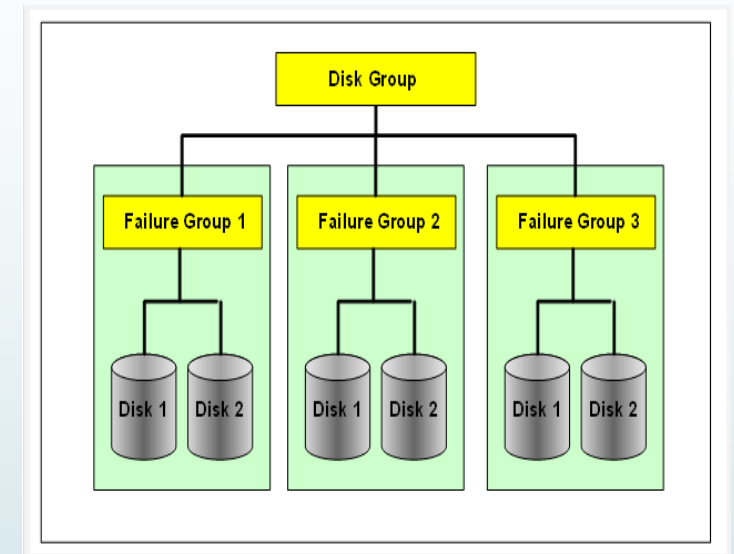
- Oracle ASM instance mountaju disk grupe koji ASM datoteke čine dostupnim instancama baza podataka
- Oracle ASM instanca i instanca baze podataka zahtjevaju dijeljeni pristup disku u disk grupi
- Upravlja metapodacima disk grupe i pruža informacije o izgledu datoteke instancama baza podataka
- Oracle Clusterware-a – jedna ASM instanca za svaki čvor klastera

Oracle ASM Cluster Configuration with Oracle RAC



Oracle ASM instance/grupe

- Disk grupa se sastoji od više diskova i osnovni je objekat kojim upravlja Oracle ASM instanca
- DATA - (najveća) koja sadrži prostor u koji se smještaju tablespace-i baze odnosno sama baza
- RECO – namijenjena za smještanje archive logova i eventualno backup-a
- REDO – sadrži redo log logove
- FLASHDATA – moguće smještanje podataka baze za brže čitanje



```

ASMCMD> lsdg
State      Type      Rebal    Sector  Block    AU       Total MB   Free MB   Req_mir_free MB   Usable_file_MB   Offline_disks   Voting_files   Name
MOUNTED    NORMAL    N        512     4096    4194304  52428736  17811724   3276796           7267464          0               Y             DATA/
MOUNTED    NORMAL    N        512     4096    4194304  1526208   440264     381552            29356           0               N             FLASH/
MOUNTED    NORMAL    N        512     4096    4194304  8618304   3370564     538644           1415960         0               N             RECO/
MOUNTED    HIGH     N        512     4096    4194304  763120    205936     381560            -58541          0               N             REDO/
  
```

Oracle ASM instance/grupe

Zauzetost diskovnog prostora

Host: [redacted].pbih.ba > Automatic Storage Management: +ASM1_[redacted].pbih.ba Logged in as SYS / SYSASM

Disk Groups


Performance **Disk Groups** Configuration Users ASM Cluster File System


Create Mount All Dismount A

Mount Dismount Rebalance Check Scrub Delete

Select All | Select None

Select	Name▲	State	Redundancy	Size (GB)	Used (GB)	Used (%)	Usable Free (GB)	Member Disks
<input type="checkbox"/>	DATA	MOUNTED	NORMAL	51,199.94	50,565.67	<div style="width: 98.76%;"></div> 98.76	-1,282.87	16
<input type="checkbox"/>	FLASH	MOUNTED	NORMAL	1,490.44	1,400.49	<div style="width: 93.97%;"></div> 93.97	-141.33	4
<input type="checkbox"/>	RECO	MOUNTED	NORMAL	8,416.31	5,124.72	<div style="width: 60.89%;"></div> 60.89	1,382.79	16
<input type="checkbox"/>	REDO	MOUNTED	HIGH	745.23	544.12	<div style="width: 73.01%;"></div> 73.01	-57.17	4

 **TIP** The usable free space specifies the amount of space that can be safely used for data. A value above zero means that redundancy can be properly restored after a disk failure.

 **TIP** Mount All and Dismount All operation will only mount and dismount the disk groups specified in the Auto Mount Disk Groups parameter.

Problemi na koje smo naišli:

- Otkazivanje memorije, diska na primarnom site-u (također i na DR site-u, istovremeno)
- Automatsko povećavanje prostora disk grupe -> veličina disk grupe premašila max. dozvoljenu veličinu za rebalans
- Rekreiranje disk grupe i disable-anje automatskog povećanja prostora nad istom

```
acfsutil size -a 0 -x 16T /u02/app/oracle/oradata/datastore
```

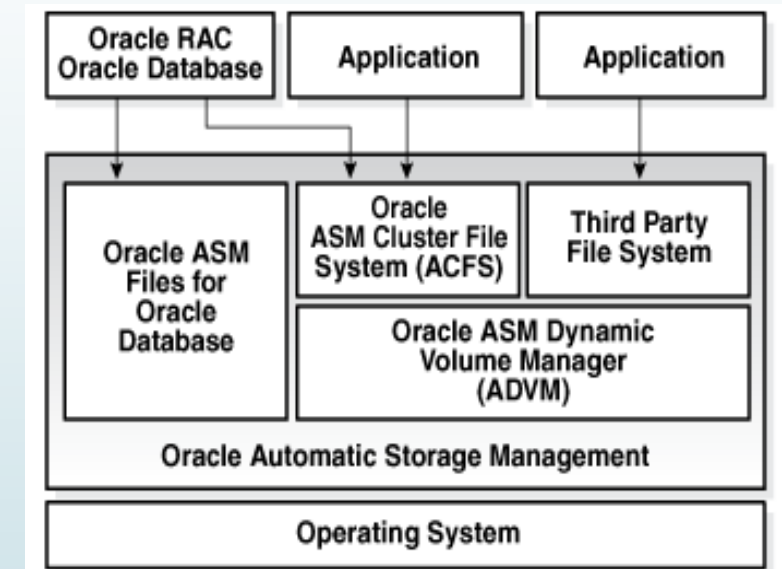
```
-a 0   disable automatskog reize-a
```

```
-x 16T za svaki slucaj maksimalna veličina dozvoljena
```

- Patch-iranje ODA-e X5-2 platforme na veću verziju

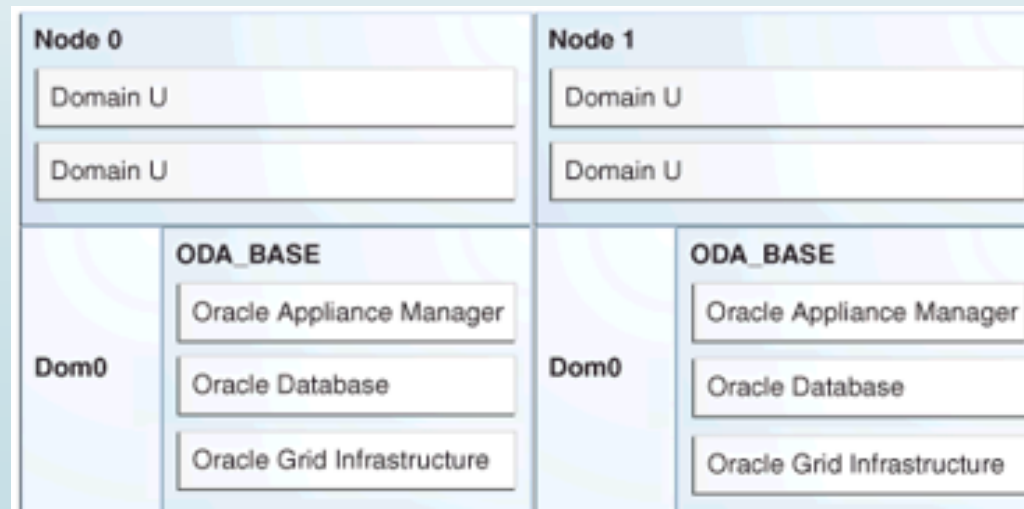
Oracle ACFS

- ACFS je predstavljen sa verzijom 11gR2
- Omogućava skladištenje aplikativnih podataka, Oracle Cluster Registry (OCR), Oracle izvršnih file-ova (binaries), sve unutar ASM-a
- ACFS je nadograđen na ASM i pruža sve beneficije ASM-a dodajući na to dodatne beneficije samog ACFS-a
- Omogućava da se i specifični ne Oracle podaci mogu pohraniti unutar cluster file sistema
- Oracle ACFS file sistem se kreira iz DATA disk grupe (za svaku bazu podataka za pohranu podataka)
- Oracle ACFS se kreira iz RECO disk grupe (za Redo i Fast Recovery area za sve baze podataka)

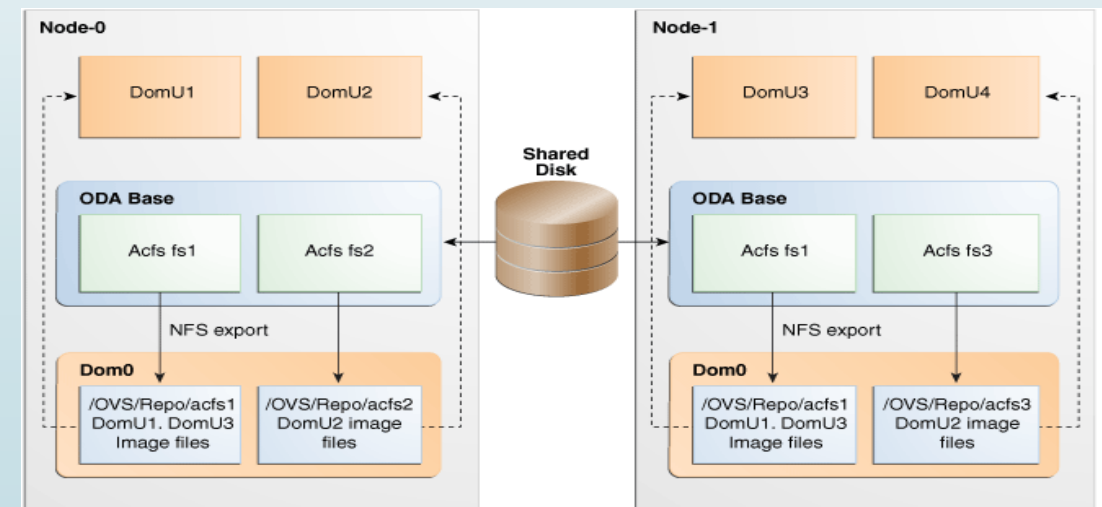


ODA VIRTUALIZIRANA PLATFORMA

- Dvije domene na svakom serverskom čvoru – Dom0 i ODA_BASE (poznat kao Dom1)
- Moguće korištenje CPU core-ova nedodijeljenih Oracle bazi podataka (ODA_BASE) virtualnim mašinama – korisnička domena ili domena (Dom)
- Oracle preporučuje korištenje zajedničkih spremišta (repositories) s virtualiziranom platformom ODA-e za visoku dostupnost

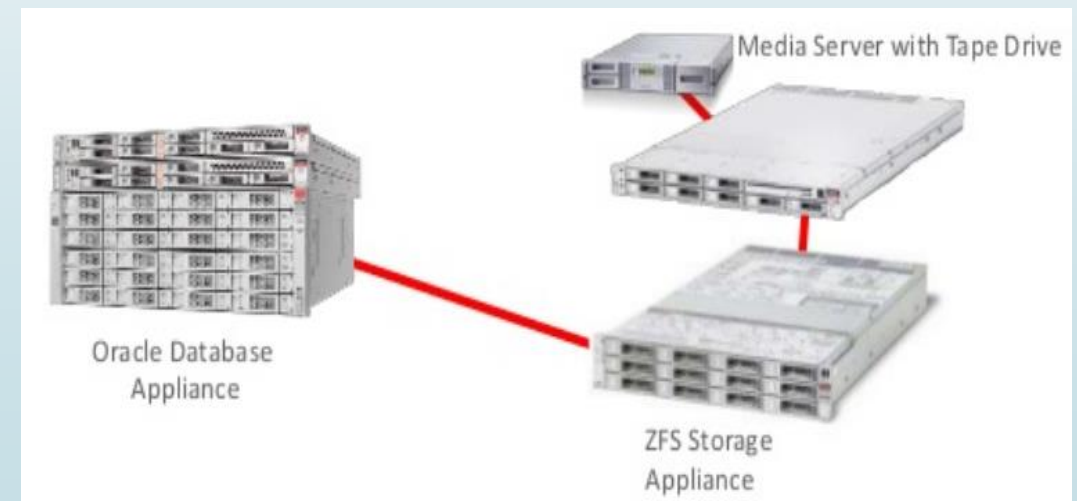


- Moguće stvaranje jednog ili više dijeljenih repositorija za pohranu datoteka VMs
- Mogućnost prelaska sa jednog čvora na drugi u slučaju kvara čvora (failover)
- Virtualni disk na dijeljenom repositoriju – dodatne skladišne opcije za VMs
- Ukoliko ne postoji potreba, storage se može odvojiti
- Virtualni diskovi se mogu dijeliti sa više VMs koje se nalaze na istom dijeljenom repositoriju



ORACLE ZFS STORAGE APPLIANCE

- Storage za višeprotokolska okruženja – ekstremne performanse, vrhunsku efikasnost za VM i okruženja u cloud-u
- Sadrži set alata za upravljanje i analitiku koji omogućava administratorima da brzo i nekomplikovano upravljaju i rješavaju probleme sa storage-om
- Mogućnost kreiranja snimaka (snapshot-a), klonova, provizioniranja, različiti algoritmi za kompresiju, replikacija preko browser-a ili CLI-a
- Mogućnost proširenja storage prostora za baze podataka koje se nalaze na ODA-i
- NFS (Network file system) protokol – prostor sa ZFS-a se dodaje na ODA uređaj
- dNFS (direct NFS) database driver



External Backup / External Storage



ORACLE ZFS STORAGE APPLIANCE

Ukupan diskovni prostor: $\approx 92,5\text{TB}$

Sun ORACLE ORACLE ZFS STORAGE ZS3-2 Super-User@ [REDACTED] LOGOUT HELP

Configuration **Maintenance** Shares Status Analytics

HARDWARE SYSTEM PROBLEMS LOGS WORKFLOWS

 [Show Details](#) ⊙ [REDACTED] ⚙️ ⏻

Manufacturer Oracle System 1.64TB (2 disks)
Model Oracle ZFS Storage ZS3-2 Data -
Serial 1537NM200W Cache -
Processors 2x2.10GHz Intel(r) Xeon(r) CPU E5-2658 0 @ 2.10GHz AES-NI Log -
Memory 256GB Total 1.64TB (2 disks)

Disk Shelves

NAME	MFR/MODEL	RPM	DATA	CACHE	LOG	PATHS	
⊙ 1537NM T00P	Oracle/Oracle Storage DE2-24C	7200	87.3TB	-	-	2	⚙️
⊙ 1537NM 400K	Oracle/Oracle Storage DE2-24P	10000	5.46TB	-	745GB	2	⚙️

ORACLE ZFS STORAGE APPLIANCE

ZFS particije - popunjenost

```

/export/backup/daily
15T 14T 1.2T 93% /mnt/zfs_storage/backup/daily
/export/backup/weekly
15T 13T 2.6T 83% /mnt/zfs_storage/backup/weekly
/export/oradata
15T 1.9T 14T 13% /mnt/zfs_storage/oradata
/export/redo_10k
200G 6.1G 194G 4% /mnt/zfs_storage/redo_10k
/export/undo_10k
1000G 96G 905G 10% /mnt/zfs_storage/undo_10k

```

Particije za backup baza podataka
(slobodno oko 3,8TB)

Particija za baze podataka
(slobodno oko 14TB)



Q & A

Rovinj, Oktobar 2019