

Sun Cluster for ORACLE Deployments

Franz Haberhauer

Technical Director

Chief Technologist Client Solutions Germany

Sun Microsystems GmbH

Sun/Oracle Relationship



Does Oracle 10g RAC need a cluster?

- Yes, it does!
- It now brings its own, called OCW/CRS
 - > Oracle 9i RAC did not have a cluster (on most OEs)
- Which offers:
 - > Virtual IPs
 - > Agent APIs
 - > Cluster membership
 - > ASM (automatic storage management)

What does SC3.x provide in addition

- SC3.x complements CRS/OCW - It is not a replacement
- CRS/OCW is well integrated with SC3.x
 - > No different “opinions” on cluster membership
- SC3.x is integrated into the Solaris kernel
 - > Gives fast and easy access to error conditions
- SC3.x offers RAC manageability features for Oracle 9i RAC
 - > Start, Stop, Monitor of RAC instances
 - > Monitoring of RAC shared storage
 - > Easy configuration of RAC parameters (UDLM, timeouts,)
- Improved integration of 10g RAC to come in the near future
- Rich set of productized agents
- Rich portfolio of agent toolkits and APIs

Solid heartbeat implementation

- SC3.x implements its heartbeat in “interrupt context”
- SC3.x heartbeat is NOT subject to scheduling problems due to high load or resource starvation!
- User land implementations are subject to false (positive) timeouts
- False positive timeouts lead to reconfiguration, which means service interruptions

Failure Fencing

- SC3.x has a very strong failure fencing implementation
 - > by design
- In case of nodes' deaths or split brain SC3.x
 - > Reconfigures and creates new (sub) cluster
 - > Sophisticated quorum mechanism
 - > Prevents access to shared storage from all nodes not part of the new cluster
 - > Using SCSI reservations (this is a safe method)
 - > Guarantees no data corruption through non-cluster nodes accessing shared data

HA interconnect

- SC3.x provides an HA IP address on the interconnect(s)
 - > up to 6 links
 - > Ethernet or Infiniband
- Should one of the interconnects fail, its failure is masked transparently to the application, and traffic is routed over the remaining interconnects
- SC3.x does application traffic striping
 - > Uses all available cluster interconnects
 - > All Oracle cache fusion traffic is striped over all of the available interconnects
 - > Striping does increase bandwidth
 - > if CPU power permits
- Cluster interconnects may also be used by other applications 7

Unique path names

- Parallel applications often prefer (or even require) to have all shared storage path names for a given lun to be identical
- Not only in heterogeneous environments this can be difficult, as HBAs (number, types, order, ...) may be different
- SC3.x provides a unique path name (didx) for a given shared lun throughout a cluster
 - > Which makes clustering of heterogeneous nodes very simple

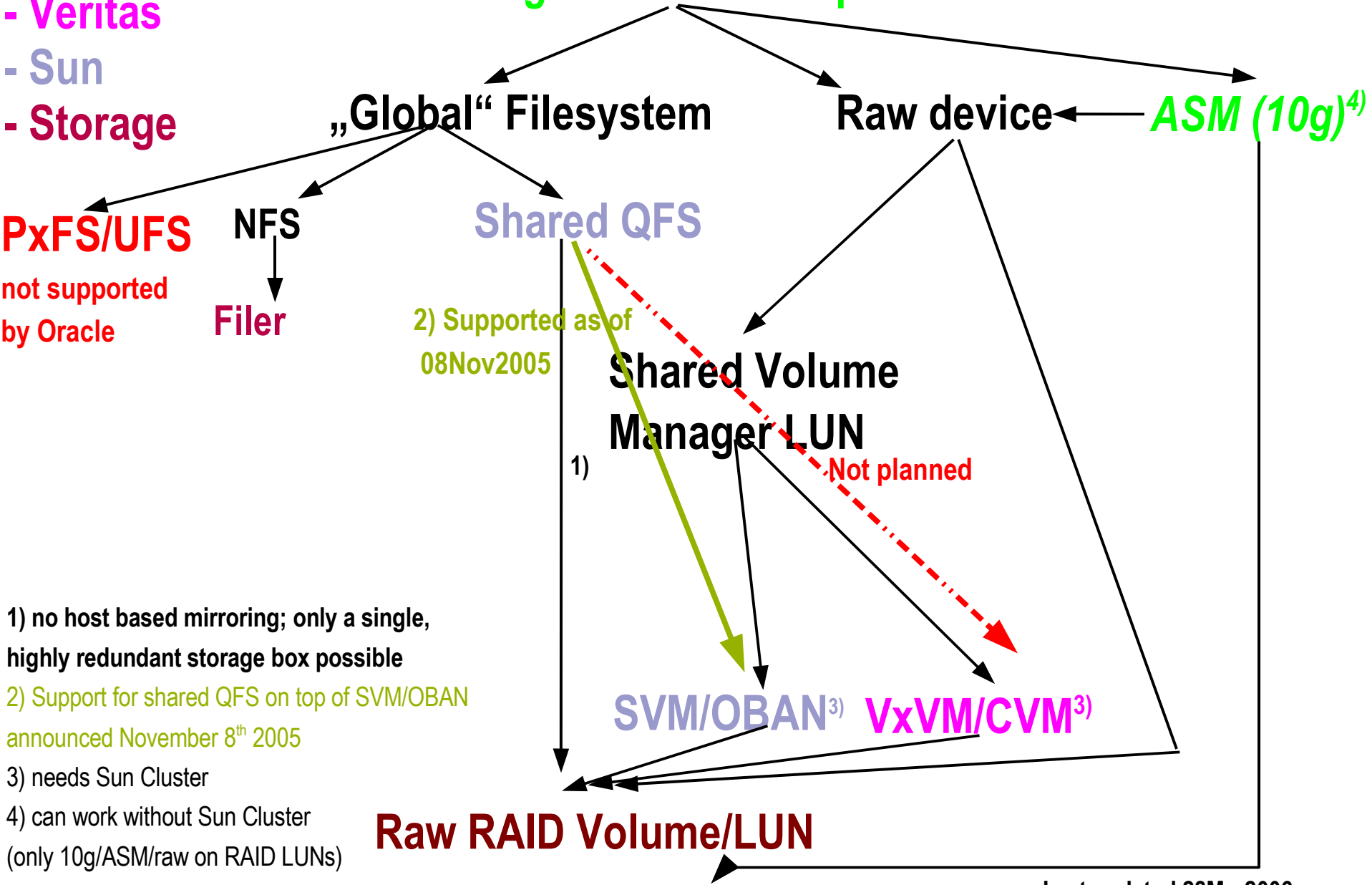
Cluster File System for Oracle RAC

- SC3.x provides support for shared QFS as a clustered filesystems that can be used to store Oracle RAC tablespaces
- Shared QFS provides direct access from all the cluster nodes to the shared storage
- Global filesystems are often asked for by system administrators – for ease of use reasons
- Shared QFS works well with Solaris Volume Manager (SVM)

- Concept
- Oracle
- Veritas
- Sun
- Storage

Oracle RAC Deployment Options

10g⁴/9i RAC tablespaces



Sun Cluster Advanced Edition For Oracle RAC



Sun Cluster for ORACLE Deployments

Franz.Haberhauer@sun.com

<http://blogs.sun.com/FranzHaberhauer>