

Tablespace copy met Dataguard

Goede oplossing voor creëren testdata

In een groot Oracle-database environment (~ 500 GB) is er veelal behoefte om over actuele testdata te beschikken. De meeste bedrijven gebruiken hiertoe een kopie van productiedata die met enige regelmaat of on demand kan worden uitgevoerd. Enkele grote applicaties voorzien in tools om dit uit te voeren, deze zijn echter niet altijd het meest effectief en kunnen bij grotere hoeveelheden data langdurig lopen. Een dergelijke datacopy kan ook vanuit de database uitgevoerd worden.

Oracle heeft hiervoor o.a. de volgende 'copy' mogelijkheden :

- Backup / Restore;
- Export en Import / Datapump;
- Database links en SELECT INTO statements;
- Tablespace copy;
- Tablespace copy in combinatie met Dataguard.

Deze mogelijkheden hebben allemaal hun voor- en nadelen. In dit artikel wordt de laatste mogelijkheid besproken, omdat deze niet echt bekend is. Om het een en ander duidelijk te maken wordt eerst kort de standaard tablespace copy besproken.

Tablespace copy zonder Oracle Dataguard

Bij een normale table space copy wordt er gebruik gemaakt van een export van metadata van de tablespace én een fysieke kopie van de datafiles waaruit deze tablespace opgebouwd is. Deze export vindt plaats op de productiedatabase. Uitgangspunt is dat normaal gesproken de productiedatabase en de testdatabase niet op een en dezelfde host geïnstalleerd zijn.

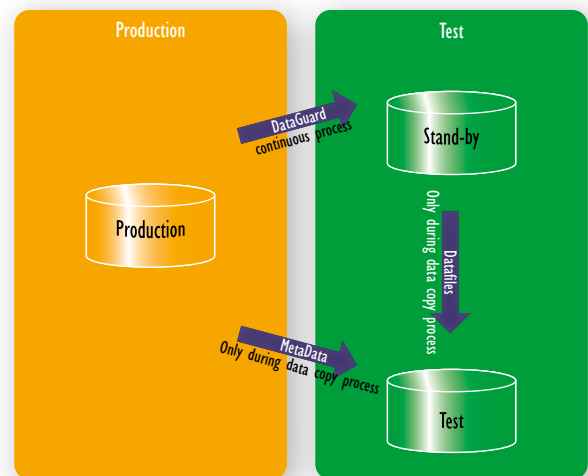
Bij een export van metadata en het kopiëren van de datafiles moeten de tablespaces die gekopieerd worden vooraf op read-only gezet worden op de productiedatabase.

De bovengenoemde metadata en datafiles moeten via een netwerk van de productie host naar de test host gekopieerd worden en geïmporteerd worden in de testdatabase. Voor kleine

omgevingen is dit nog wel te doen ; voor grotere omgevingen (> 100 GB) duurt dit natuurlijk veel langer met als bijkomend nadeel de extra belasting van het netwerk. Vervolgens moeten de tablespaces geïmporteerd worden met behulp van de gekopieerde metadata en datafiles in de testdatabase.

De nadelen van een standaard tablespace copy van een productie naar een test host zijn:

De productie database, waar de tablespaces van gekopieerd moeten worden, moeten op read-only staan tijdens de kopie waardoor deze database niet volledig gebruikt kan worden voor de volledige duur van het proces; Het kopiëren van datafiles van de productie host naar de test host kost veel tijd en heeft een hoge netwerkbelasting.



Tablespace copy met Oracle Dataguard

Voor de kopie van de datafiles kan Oracle Data Guard een oplossing bieden. Oracle Data Guard is een standaard Oracle-oplossing voor Oracle-databases in een disaster-recovery scenario. In dit voorbeeld wordt gebruik gemaakt van een fysieke stand-by database. De stand-by database wordt bijgewerkt door archived logs van de productie (of primaire) database. In

deze archived logs staan alle transacties die op de primaire database zijn uitgevoerd. De stand-by database staat in recovery mode en verwerkt deze archived logs die van de primaire database overgezonden worden. Dit kan direct zijn, maar er kan ook een delay op gezet worden met als bijkomend voordeel dat gegevens uit het verleden (bijvoorbeeld maximaal 1 uur) eenvoudig kunnen worden teruggehaald of bekeken kunnen worden.

Omdat de stand-by database exact dezelfde structuur heeft als de productiedatabase, het is immers een fysieke stand-by database, kan ook gebruik gemaakt worden van de datafiles van de stand-by Oracle-database voor de tablespace copy.

Het proces

In onderstaande beschrijving wordt uitgegaan van twee hosts die via een netwerk met elkaar verbonden zijn.

Productie host : waar de productie database op draait;

Test host : waar de test database en de stand-by database op draait.

Voor de tablespace copy wordt gebruik gemaakt van de metadata van de te kopiëren tablespaces van de productiedatabase. Van de metadata van de stand-by database kan geen gebruik gemaakt worden, omdat deze alleen in recovery- en read-only modus kan staan. Voor een export van metadata van tablespaces moet de database in read-write modus (standaard) staan en kan dus alleen gebruik gemaakt worden van de metadata van de tablespaces van de productie database.

Voor de export van de metadata moeten de tablespaces op de productiedatabase die gekopieerd moeten worden wel op read-only gezet worden. Dit duurt echter maar enkele minuten (een metadata export voor een 500Gb database kan binnen 5 minuten).

Aangezien Dataguard actief is, worden de tablespaces op de stand-by database ook read-only gezet en dus ook de datafiles van deze tablespaces. Zodra de datafiles op de stand-by database read-only staan wordt deze tijdelijk uitgezet (shutdown), zodat er niets meer wijzigt aan de datafiles van de stand-by database.

Alleen de metadata van de tablespaces van de productiedatabase moet geëxporteerd en vervolgens gekopieerd worden van de productie host naar de test host. De datafiles van de te kopiëren tablespaces van de productie host staan immers al op de test host middels de stand-by database.

De datafiles op de test host hoeven alleen gekopieerd te worden van de stand-by locatie naar de test locatie. Men hoeft dus geen datafiles meer te kopiëren over het netwerk, kopiëren op dezelfde machine, de test host, gaat veel sneller.

Zodra de datafiles van de stand-by database naar de plaats voor de testdatabase zijn gekopieerd zijn, kan er met behulp van metadata van de productiedatabase een import gedaan worden van de tablespaces in de testdatabase.

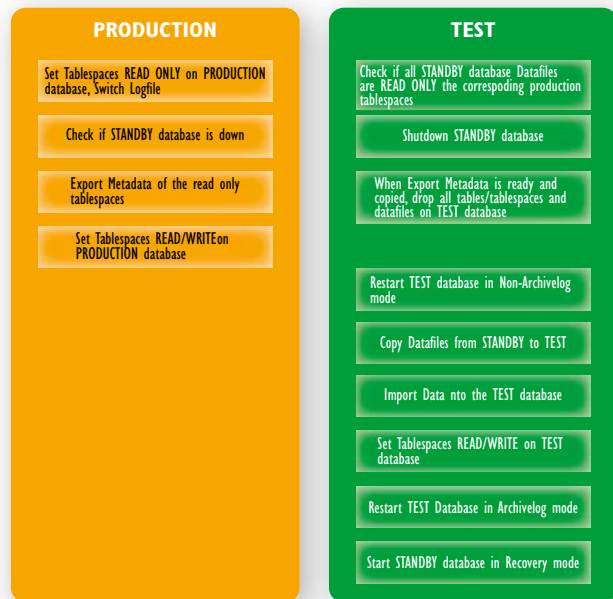
De stand-by database kan na de import weer gestart worden in recovery modus.

Voordelen

De voordelen van een tablespace copy in combinatie met Oracle Dataguard zijn:

Geringe downtime van Productie-omgeving. Productiedatabase tablespaces hoeven maar zeer korte tijd op read-only gezet te worden voor het exporteren van de metadata.

Geen netwerk belasting, er worden geen datafiles over het netwerk gekopieerd van productie host naar de test host. Deze staan er immers al bij de stand-by database.



Figuur 2: Het kopieproces schematisch weergegeven.

In een praktijkvoorbeeld wordt circa 500 GB aan datafiles gekopieerd. Totaal duurt dit het gehele proces 2,5 uur. Voorheen duurde dit minimaal 10 uur was gedurende de tijd de productie omgeving niet beschikbaar omdat de productie data die gehele tijd op read-only stond. De individuele stappen van het gehele proces zoals in dit artikel is beschreven is weergegeven in Figuur 2.

Conclusie

Dataguard is niet alleen een goede oplossing voor disaster recovery maar zeker ook voor replicatie van productiedata naar testdata.



Paul Vermeulen is Senior Oracle DBA bij Quistor Enterprises.