

Oracle GoldenGate – DB Migration auf eine Exadata Database Machine mit Fallbackszenario

Konrad Häfeli
Trivadis AG
CH-3014 Bern

Schlüsselworte:

Oracle GoldenGate, Oracle Database Machine, Exadata, Replikation, Migration

Einleitung

Migrationen sind immer mit Risiken verbunden, die Risikoanalyse und die Minimierung der Risiken ist ein Hauptteil der Migrationsvorbereitung und widerspiegelt sich in den jeweiligen Migrationskonzepten. Je mehr Unbekannte im System sind, desto grösser sind die Risiken und demzufolge auch die Massnahmen die zur deren Minimierung getroffen werden müssen. Es gibt keine Veränderung ohne Problempotential, aber der Lebenszyklus von Hard- und Software sowie des Geschäftes allgemein bedingt, dass wir uns mit Veränderungen auseinandersetzen.

Testen der Migrationsabläufe ist ein Muss, damit der funktionelle Ablauf und das zeitliche Verhalten bekannt sind. Automatismen sichern das korrekte, reproduzierbare Vorgehen. Meistens bleiben aber immer noch Unsicherheiten betreffend dem Verhalten der Neuen Umgebung, wenn die wirkliche Last, die realen Anwender mit echten Daten auf die Systeme zugreifen. Damit solche Risiken abgedeckt werden können müssen sogenannte Fallbackszenarien definiert werden, mit denen man auch nach dem Start des produktiven Betriebes wieder auf die alte, bewährte Umgebung zurückgehen kann. Diese Szenarien sind meist sehr aufwändig, sind doch nicht nur Infrastrukturteile involviert, sondern auch produktive Daten.

Eine weitere grosse Herausforderung ist die Minimierung der Ausfallzeit während der Migration. Sind da hohe Anforderungen, sprich nur kleine Unterbrüche möglich, werden gewisse Szenarien zum vornherein wieder ausgeschlossen.

Oracle hat mit der Akquisition von GoldenGate im letzten Jahr ein Produkt erworben, das für beide Fälle interessant ist, sowohl die Minimierung der Ausfallzeit, wie auch die Möglichkeit eines Fallbacks auf die alte Umgebung. Beide Male werden Daten repliziert, zuerst von der alten Produktion auf das neue System, danach vom neuen produktiven System auf das alte System.

In diesem Bericht wird das Migrationskonzept erklärt, die Komponente GoldenGate beschrieben und die praktischen Erfahrungen (positive und negative) damit aufgezeigt.

Systemkomponenten

Das Projektsetup für die Migrations- und Fallbackphase beinhaltet folgende Komponenten:

- Quellsystem (RAC Datenbank auf Sun E2900)
- Zielsystem (RAC Datenbank auf Quarter Rack, Exadata Database Machine)
- Oracle GoldenGate für die Replikation der Daten
- DBFS (Datenbank Filesystem) für die Bereitstellung der Replikationsdaten
- Real Application Testing für die Messung der Performance unter produktiver Last

Oracle Exadata Database Machine Übersicht

Oracle steigt mit Sun Microsystems ins Hardware Business ein. Die verschiedenen Versionen der Exadata Database Machine wurden alle mit den besten Performanceattributen versehen. Im Gegensatz zur Version 1 der Database Machine (in Zusammenarbeit mit HP) adressiert die Sun Oracle Database Machine nicht nur Performanceprobleme im Data Warehousing (DW) sondern auch im Online Transaction Processing (OLTP), sowie im gemischten Betrieb.

Das Geheimnis dieser Appliance steckt in der Verbindung zwischen der Sun Hardware und der Oracle Datenbank. Eine grosse Bandbreite verbindet diese beiden Elemente, welche einen enormen Datendurchsatz erlaubt. Für DW optimal werden die SQL-Abfragen auf dem intelligenten Storage vorverarbeitet, sodass nur ein Bruchteil der Datenmenge vom Storage zu den Datenbankinstanzen transferiert werden muss. Die Lösung für Transaktionsverarbeitung liegt im Smart Flash Cache der auf der Exadata Storagezelle implementiert ist. Durch den Einsatz von intelligent verwalteten PCIe Flash Cards werden noch bessere I/O Werte möglich als wenn die Disks nur durch Flash Disks ersetzt würden. Auch die Columnar Compression Funktionalität, welche nur auf Exadata Storage verfügbar ist, trägt durch massive Reduktion der Datenmenge ein grösser Anteil zum Performancegewinn bei.

Der Lebenszyklus der Exadata Database Machine hat folgende Ausprägungen hervorgebracht:

- V1 mit HP (Ende 2008, 2009) für DWH
- V2 mit Sun (Herbst 2009, OOW¹) DWH und OLTP
 - 8 Server mit je 2 Sockets à **4 Cores**
- X2-2 mit Sun (Herbst 2010, OOW)
 - 8 Server mit je 2 Sockets à **6 Cores**
- X2-8 mit Sun (Herbst 2010, OOW)
 - 2 Server mit je 8 Sockets à **8 Cores**



¹ Oracle Open World Konferenz in San Francisco

Die V2 und X2-2 sind als Quarter-, Half- und Full-Rack Version lieferbar, die X2-8 wird nur als Full-Rack geliefert, da sie nur aus 2 Server besteht die durch eine Halbierung keine Redundanz mehr hätten.

Die an diesem Migrations-Projekt eingesetzte Database Machine ist ein V2 Quarter-Rack mit 2 Database Server und 3 Storage Server, welche in sich ein von Oracle vorkonfiguriertes, ausbalanciertes System darstellen.

Oracle GoldenGate Übersicht

Die Funktionalität von Oracle GoldenGate ist sehr umfangreich. Einsatzkonzepte, Kommandosyntax und Automatisierungsmöglichkeiten sind ebenfalls sehr vielfältig. In diesem Artikel wird nur eine grobe Übersicht anhand eines Projekteinsatzes gegeben, was noch keine best practice widerspiegelt. In naher Zukunft wird der Autor aber die gemachten Projekterfahrungen zusammenfassen und konkretere Umsetzungsempfehlungen geben können.

Oracle positioniert die Akquisition vom letzten Jahr im Bereich Middleware und belegt diese mit den Thematiken von

- Echtzeit Daten Integration
- Kontinuierliche Datenverfügbarkeit
- Kontinuierliche Datensynchronisation

dies über heterogene Systeme hinweg. D.h. also nicht nur von Oracle nach Oracle über unterschiedliche Versionen und Plattformen, sondern auch von anderen Produkten auf Oracle. Die Datentransfers sind in beide Richtungen möglich, klar dass Oracle die Migrationsvariante AUF die Oracle Datenbank dem Szenario WEG von der Oracle Datenbank vorzieht. Die Funktionalität ist dem Portfolio der Oracle Data Integration Solution zugeordnet:

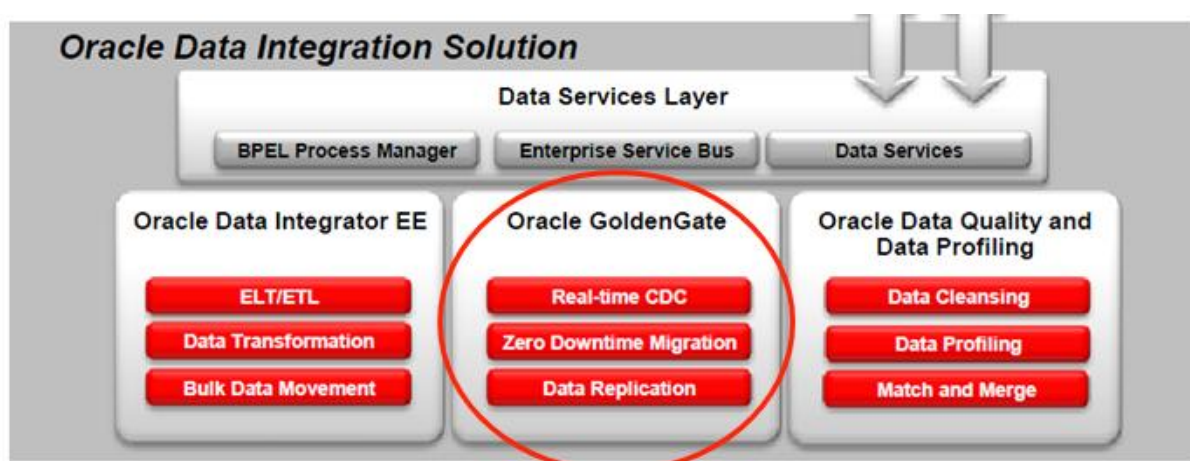


Abb. 1: Oracle Data Integration Solution Portfolio mit GoldenGate (Quelle Oracle)

Die Replikation von Daten kann in verschiedenen Topologien gemacht werden:

- one-to-many
- many-to-many
- cascading
- bidirectional

Diese Topologien erlauben dann nebst anderen Anwendungen auch den Einsatz bei (beinahe) „zero-downtime-upgrade/migration“ Szenarien.

Oracle GoldenGate Architektur

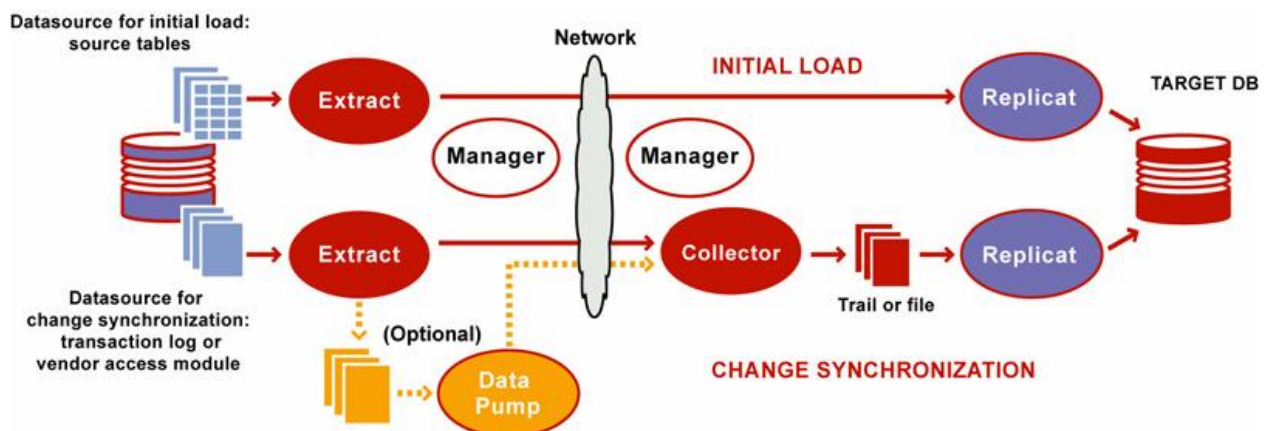


Abb. 2: Architektur Oracle GoldenGate (Quelle Oracle)

GoldenGate bestehe aus folgenden Komponenten:

- **Extract:**
Dieser Prozess läuft auf dem Quell-System und liest die Transaktions-Daten aus den online redologs (oder archived redologs). Nebst den inserts, updates und deletes können auch die DDL Kommandos extrahiert (capture) werden. Dies macht die Synchronisation einer Zieltabellebank möglich. Extract kann ebenfalls eingesetzt werden um das initiale Laden (initial load) des Zielsystems zu machen. Es sind mehrere Extract-Prozesse möglich die auf verschiedene Objekte angesetzt werden können.
- **Data Pump:**
Dieser Prozess läuft ebenfalls auf dem Quell-System und kann eingesetzt werden um die lokal geschriebenen Extract-Daten (local trail) auf ein Remote-System zu verschieben (remote trail). Dies ermöglicht die Entkopplung der zwei Systeme, da der local und remote trail als File-Buffer dienen können. Die Data Pump kann die Daten ungeändert durchreichen (pass-through-mode) oder diese auch filtern oder konvertieren.

- **Replicat:**
Dieser Prozess läuft auf dem Ziel-System, liest Änderungen und DDLs aus den trail-Files und appliziert diese. Es können ebenfalls parallele Replicat-Prozesse gestartet werden die auf verschiedene Objekte konfiguriert sind. Das Applizieren kann zeitverzögert werden, damit eventuelle logische Fehler nicht gleichzeitig repliziert werden.
- **Trail extract files:**
Die Trail-Files beinhalten die Datenänderung, werden von den Extract Prozessen geschrieben und von den Data Pump und Replicat Prozessen gelesen. Die Files werden sequentiell mit sechs Ziffern im Namen durchnummeriert und haben default eine Grösse von 10MB. Der Inhalt wird proprietär geschrieben (universal data format) und ist optimiert auf maximalen I/O Durchsatz.
- **Checkpoints:**
GoldenGate speichert die Schreib- und Lesepointer in Checkpoint-Files. Dies ermöglicht im Fehlerfall das Richtige aufsetzen der Prozesse. Diese proprietäre Technologie garantiert die Meldungsübertragung über das Netzwerk hinweg.
- **Manager:**
Auf beiden Seiten (Quelle und Ziel) läuft ein Managerprozess welcher die konfigurierten Prozesse überwacht und im Fehlerfall automatisch restartet. Er ist ebenfalls zuständig für das Logging und Reporting von Fehlern und Events.
- **Collector:**
Der Collector liest via Netzwerk die Daten von den Extract Prozessen und schreibt sie in die Trail-Files. Er läuft auf dem Zielsystem und kann statisch (manuell) oder dynamisch (via manager) gestartet werden. Bevorzugt wird eine eins-zu-eins Beziehung zu den Extract-Prozessen um eine optimale Performance zu erreichen.

Innerhalb dieser Komponenten unterhält GoldenGate eine *Comit Sequence Number* (CSN), welche eine platformunabhängige Nummer für die SCN von Oracle darstellt. Dadurch wird die Transaktionskonsistenz und Datenintegrität sichergestellt. Die CSN stellt einen bestimmten Zeitpunkt im Leben einer Datenbank dar, welcher sich in der Replikation widerspiegelt.

Oracle GoldenGate Konfiguration

GoldenGate hat standardisierte Verzeichnisnamen in welchen die verschiedenen Files abgelegt werden. Diese liegen unterhalb des Home-Verzeichnisses von GoldenGate und werden mit dem Befehl *create subdirs* innerhalb des Programms *ggsci*² erstellt:

² GoldenGate Software Command Interface, ist das Command Line Interface (CLI) zwischen Anwender und Komponenten

- `dirdat`: beinhalte die Trailfiles
- `dirchk`: beinhaltet die Checkpointfiles
- `dirpcs`: beinhaltet die Statusfiles, die File-Extensions zeigen den Bezug zum jeweiligen Prozess (`.pce` für Extract, `pcr` für Replicat, `pcm` für Manager)
- `dirprm`: beinhaltet die Parameterfiles
- `dirrpt`: beinhaltet die Reportfiles
- `dirsql`: beinhaltet die Sql-Files für die Erstellung von Triggern und Logtabellen
- `dirtmp`: beinhaltet die temporären Files welchen geschrieben werden wenn grosse Transaktionen nicht mehr im Memory gehalten werden können

Die Erstellung der Directory-Struktur, der Parameter-Files und die Manipulation, sowie das Monitoring der Umgebung kann mit einem Dienstprogramm (CLI) gemacht werden. Das `ggsci` liegt im Versions-Verzeichnis der GoldenGate Installation, welche durch das Entpacken des GoldenGate Tarballs (Download via <http://edelivery.oracle.com>) an der gewünschten Stelle erstellt wird.

Folgende Parametersettings unterstützen das Arbeiten:

```
export PATH=/u00/app/goldengate/10.4:$PATH
export LD_LIBRARY_PATH=/u00/app/goldengate/10.4:$LD_LIBRARY_PATH
```

GoldenGate benötigt in der Quell- und Ziel-Datenbank einen Useraccount mit entsprechenden Privilegien. Ebenfalls muss supplemental logging eingeschalten werden um die nötigen Informationen in die Redostrukturen schreiben zu lassen.

```
ALTER DATABASE ADD SUPPLEMENTAL LOG DATA;
ALTER SYSTEM SWITCH LOGFILE;
```

Die Konfiguration der Extracts und Replicats kann entweder direkt durch das Editieren der Parameterfiles gemacht werden oder durch die Eingabe der nötigen Kommandos im `ggsci`-CLI. Das Tool hat folgende Hauptbefehle:

- `start`, `stop`, `status`, `send`, `add`, `alter`, `info`, `delete`, `stats`

Diese beziehen sich auf den Manager, sowie die Extract und Replicat Prozesse, welche jeweils Namen zugeordnet haben.

Als einfaches Beispiel sollen die Tabellen `emp` und `dept` aus dem Schema `scott` repliziert werden. Dazu werden zuerst die Parameterfiles für die Prozesse geschrieben und danach vom Managerprozess gestartet.

Manager Parameterfile:

```
--Port on which the manager listens for incoming messages
PORT 7809
--List of reserved ports for the process communication
```

```
DYNAMICPORTLIST 7900-7925
--Manager tries to start all Extract processes 15 times after failure
--within 5 min time period, and it waits 10 minutes before each attempt
AUTORESTART EXTRACT *, RETRIES 15, WAITMINUTES 10, RESETMINUTES 5
```

Extract Parameterfile:

```
EXTRACT xscott1
USERID ggs, PASSWORD *****
--report problems
DISCARDFILE /u00/app/goldengate/dirrpt/xscott1.dsc, PURGE
--trailfile directory on the local system
EXTTRAIL /u00/app/goldengate/dirdat
TABLE scott.emp;
TABLE scott.dept;
```

Data Pump Parameterfile:

```
EXTRACT pscott1
--run in passthru mode, no connection with db is needed
PASSTHRU
--name of the target machine, and the mgr port on the target machine
RMTHOST targethost, MGRPORT 7809
--trailfile directory on the remote system
RMTTRAIL /u00/app/goldengate/dirdat
TABLE scott.emp;
TABLE scott.dept;
```

Replicat Parameterfile:

```
REPLICAT rscott1
USERID ggs, PASSWORD *****
-- source and target tables are the same
ASSUMETARGETDEFS
--report problems
DISCARDFILE /u00/app/goldengate/dirrpt/rscott1.dsc, PURGE
--trailfile directory on the local system
EXTTRAIL /u00/app/goldengate/dirdat
TABLE scott.emp;
TABLE scott.dept;
```

Auf beiden Systemen werden die Prozessgruppen hinzugefügt und über den Managerprozess gestartet.

Quellsystem:

```
$ ggsci
GGSCI> CREATE SUBDIRS
GGSCI> USERID ggs, PASSWORD *****
GGSCI> ADD EXTRACT xscott1, TRANLOG, BEGIN now
GGSCI> START EXTRACT xscott1
GGSCI> INFO EXTRACT xscott1, DETAIL
GGSCI> ADD EXTRACT pscott1, TRANLOG, BEGIN now
GGSCI> START EXTRACT pscott1
GGSCI> INFO EXTRACT pscott1, DETAIL
GGSCI> START MGR
GGSCI> STATUS ALL
```

Zielsystem:

```
$ ggsci
GGSCI> CREATE SUBDIRS
GGSCI> USERID ggs, PASSWORD *****
GGSCI> ADD REPLICAT rscott1
GGSCI> START REPLICAT rscott1
GGSCI> INFO REPLICAT rscott1, DETAIL
GGSCI> START MGR
GGSCI> STATUS ALL
```

Das automatisierte Monitoring der Umgebung benötigt noch einige Scripte auf Shell Ebene, das Kommando

```
GGSCI> INFO ALL
```

listet für alle Prozesse (Monitor, Extract, Replicat) die Stati und die LAG Informationen (Differenz der Prozessverarbeitung zum effektiven Zeitpunkt auf dem Quellsystem) auf. Das Utility LOGDUMP lässt detaillierte Analyse der Trailfiles zu und ermöglicht so die Problembehandlung auf Transaktionsebene.

Initiales Laden

GoldenGate hat verschiedene Möglichkeiten die Zieldatenbank initial zu laden um sie danach auf Transaktionsebene nachzuführen. Für einen Oracle DBA liegt die Methode mit einem RDBMS Utilities wohl am Nächsten. D.h. der Einsatz von export/import datapump (expdp, impdp) lässt sich parallel zum produktiven Betrieb machen. Folgender Ablauf muss dabei eingehalten werden:

- Konfiguration und Starten des GG Extract Prozesses
- Konsistenter Export mit expdp und FLASHBACK_SCN gesetzt (aktuelle Quell-SCN)
- Import der zeitkonsistenten Daten auf dem Zielsystem
- GG Replicat Prozess mit begin CSN auf die export FLASHBACK_SCN

Exadata (RAC) Spezifika

Hochverfügbarkeit und Recoverability des Replikationsprozesses bei GoldenGate kann man erreichen wenn die Trail- und die Checkpoint-Files auf einem für alle RAC Knoten gemeinsamen Storage liegen. Dies kann ab Oracle 11gR2 mit einem DBFS (Database Filesystem) erreicht werden. Da dieses Filesystem durch die Datenbank bereitgestellt wird, wird es auch bei einem Knotenfehler durch die anderen verfügbaren Knoten zur Verfügung gestellt und bietet somit dem Replicat Prozess der über die Oracle Clusterware mit einer Failoverkonfiguration restartet wird, die benötigten Files an.

Transaktionsdaten werden in Memorybuffern gehalten bis das Commit resp. Rollback kommt. Sind diese Buffer voll werden Sie in sogenannte Page Files (ins Directory dirtmp)

geschrieben. Da dieser Filetyp nicht auf einem DBFS liegen kann, müssen sie lokal pro Knoten gespeichert werden.

Fallback Szenario

GoldenGate mit seinem platform- und versionsunabhängigen Format kann so konfiguriert werden, dass nach einer Übernahme der neuen Exadata Platform die Extract und Replicat Prozesse in umgekehrter Richtung konfiguriert werden. Dies ermöglicht die Replizierung der Änderungen zurück auf die alte Platform, was die Möglichkeit einer Wiederaufnahme der Produktion offen lässt. Diese Konfiguration kann eine gewisse Zeit laufen bis sichergestellt ist, dass das neue System seine Arbeit zur Zufriedenheit erfüllt.

Zu beachten/Probleme

Auf dem Zielsystem sollten die Trigger und die referentielle Integrität ausgeschaltet werden (Foreign Key Constraints). Dadurch verhindert man Probleme die durch DML welche von Triggern resp. Delete cascade Aktivitäten kommen, ausgelöst werden. Diese Datenmanipulationen werden ebenfalls durch GoldenGate repliziert und würden dann zu Fehlern führen. Mittels generierten DDL Scripten, können disable- und enable-Scripte bereitgestellt werden die vor dem Start der Replikation (disable) oder nach der Datenmigration (enable) ausgeführt werden können.

Sequenzen sind auch immer ein Problem in replizierten Umgebungen. Bei einer Active/Standby Variante können aber die Sequenzen wie sie auf dem Quellsystem sind auf das Ziel übertragen werden.

Grundsätzlich muss entschieden werden ob DDLs auch repliziert werden. GoldenGate kann das, aber in einem Migrationsprojekt macht es nicht unbedingt Sinn dies zu implementieren, denn in einer Migrationsphase sollte die Datenstruktur stabil bleiben. Dies vereinfacht das Setup und den Betrieb des Systems.

Ein Problem wurde festgestellt, nachdem die Exadata die Produktion übernommen hatte. Die GoldenGate Prozesse hatten auf dem neuen System wider erwarten eine grosse CPU Last gemacht. Der Ursache konnte aus zeitlichen Gründen nicht auf den Grund gegangen werden, durch die Separierung der Prozesse auf einem Knoten und der grossen Reserve der Database Machine beeinträchtigte das den Betrieb aber nicht.

Fazit

Die Replikation von Transaktionen wurde innerhalb von Oracle Datenbanken bisher meistens mit Streams gemacht. Die Komplexität, Stabilitäts- und Performanceprobleme einer Migration mit Streams stellen die Variante mit GoldenGate aber als optimalere Lösung hin. Ohne Einschränkungen und Scriptaufwand geht aber auch das nicht. Die Einfachheit der Architektur, die Geschwindigkeit und Flexibilität lassen aber den Einsatz von GoldenGate in vielen Fällen trotz hoher Kosten in Betracht ziehen. Durch das Einsetzen von Funktionalitäten wie DBFS und Grid Infrastructure können die Änderungsdaten und Prozesse auch

hochverfügbar gemacht werden und somit eine möglichst kleine Replikationsverzögerung auch im Systemfehlerfall garantiert werden.

Mit Oracle GoldenGate kann eine Plattformmigration mit kleiner Unterbruchszeit durchgeführt werden und eine Fallbackinfrastruktur mit aktuellen Daten bereitgestellt und nachgeführt werden.

Kontaktadresse:

Konrad Häfeli

Trvadis AG

Papiermühlestrasse 73

CH-3014 Bern

Telefon: +41(0)31-928 09 60

Fax: +41(0)31-928 09 64

E-Mail konrad.haefeli@trivadis.com

Internet: www.trivadis.com