Read–Write Replication in Arla

Författare: Noora Peura

Examinator: Karl–Filip Faxén
Handledare: Fredrik Lundevall
Uppdragsgivare: NADA, KTH
Industrihandledare: Ragnar Andersson

Innehåll

Målet med detta arbete

Föreslå ett eller fler sätt att implementera skrivreplikering av volymer i det distribuerade filsystemet Arla.

Introduktion till AFS och Arla

Innehåll:

- AFS
- Servrar i AFS
- AFS–klienten
- Arla

Introduktion till AFS och Arla

AFS:

- Transparent distribuerat filsystem
- Klient−serverbaserat
- Volymbaserad administration
- Stöd för läsreplikering av volymer

Introduktion till AFS och Arla

Servrar i AFS:

- Volume Server
- Volume Location Server
- File Server
- m.fl.

- Flera av varje server
- Replikerade databaser i flera av servrarna

Introduktion till AFS och Arla

AFS−klienten (Cache Manager):

- En klient på varje arbetsstation
- Arbetar mot VL−servern och File−servern
- Cache på lokal disk
- Callbacks

Introduktion till AFS och Arla

Arla:

- Fri AFS−implementation
- Klienter för många plattformar
- Server under utveckling

Replikeringsalgoritmer

Innehåll:

- Varför replikering
- Svårigheterna i skrivreplikering
- Optimistisk eller pessimistisk replikering
- Replikering med primär nod
- Replikering med omröstning

Replikeringsalgoritmer

Varför replikering?

- Ökad tillgänglighet
- Lastbalansering

Fördelarna beror till stor del på replikeringsalgoritmen!

Replikeringsalgoritmer

Svårigheterna i skrivreplikering:

- Distribuera uppdateringar tillförlitligt
- Hålla reda på senaste versionen
- Hantering av onormala situationer

Skrivreplikering kostar alltid!

Replikeringsalgoritmer

Optimistisk eller pessimistisk replikering?

- Optimistisk replikering: skriv när du vill, ta hand om problemen när de uppstår!

- Pessimistisk replikering: se till att det inte blir problem!

Replikeringsalgoritmer

Replikering med primär nod:

- Alla skrivningar och läsningar skickas till en primär nod
- Primära noden skickar ut ändringar till sekundärerna
- Val av ny primär nod när den gamla går ner
- Ger ej lastbalansering!

Replikeringsalgoritmer

Replikering med omröstning:

- Koordinator utses
- En beslutsmässig grupp av noder måste delta i varje läs− eller skrivoperation
- Flera sätt att bilda beslutsmässiga grupper:
  - Majoritet
  - Skriv till alla/läs från en
  - m.fl.
- Fördelarna och nackdelarna beror på algoritmen

Förslag på replikeringsmetoder

Innehåll:

- Krav på replikeringsalgoritmen
- Replikeringsalgoritmer
- Asynkrona eller synkrona uppdateringar
- Återhämtning efter krascher

Förslag på replikeringsmetoder

Krav på replikeringsalgoritmen:

- AFS−kompatibel
- Konsistent filsystem
- Ökad tillgänglighet
- Minimal påverkan på prestandan

Förslag på replikeringsmetoder

Replikeringsalgoritmer:

- Primär nod
- Omröstning med majoritet
- Skriv till alla/läs från en
- Optimistisk algoritm (Available Copy)

Förslag på replikeringsmetoder

Asynkrona eller synkrona uppdateringar?

- Asynkron: skrivoperationen returnerar snabbt, men skrivningen kan ha misslyckats.
- Synkron: skrivoperationen returnerar när skrivningen har lyckats/misslyckats.

Förslag på replikeringsmetoder

Återhämtning efter kraschen:

- "Ivrig" återhämtning: synkronisera filsystemet direkt efter kraschen.
- "Lat" återhämtning: synkronisera filsystemet när det behövs.

Rekommendationer

Allmänt:

- Synkrona uppdateringar
- Ivrig återhämtning

Alternativ:

- Skriv till alla/läs från en
- Omröstning med majoritet
- Kombination av båda
- Ingen replikering