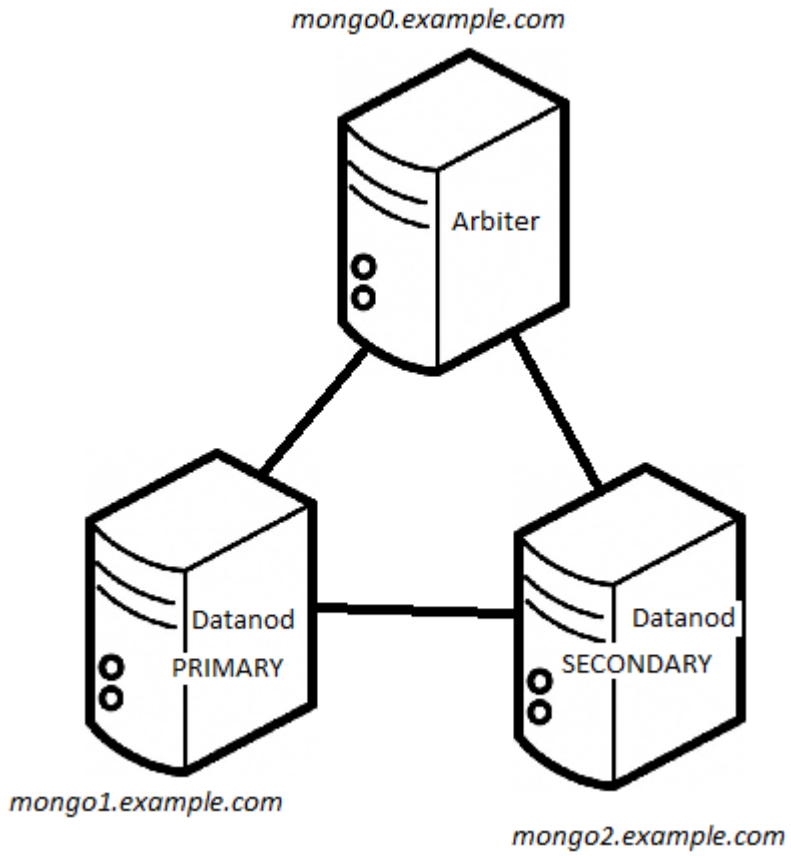
MongoDB för Säkerhetstjänster

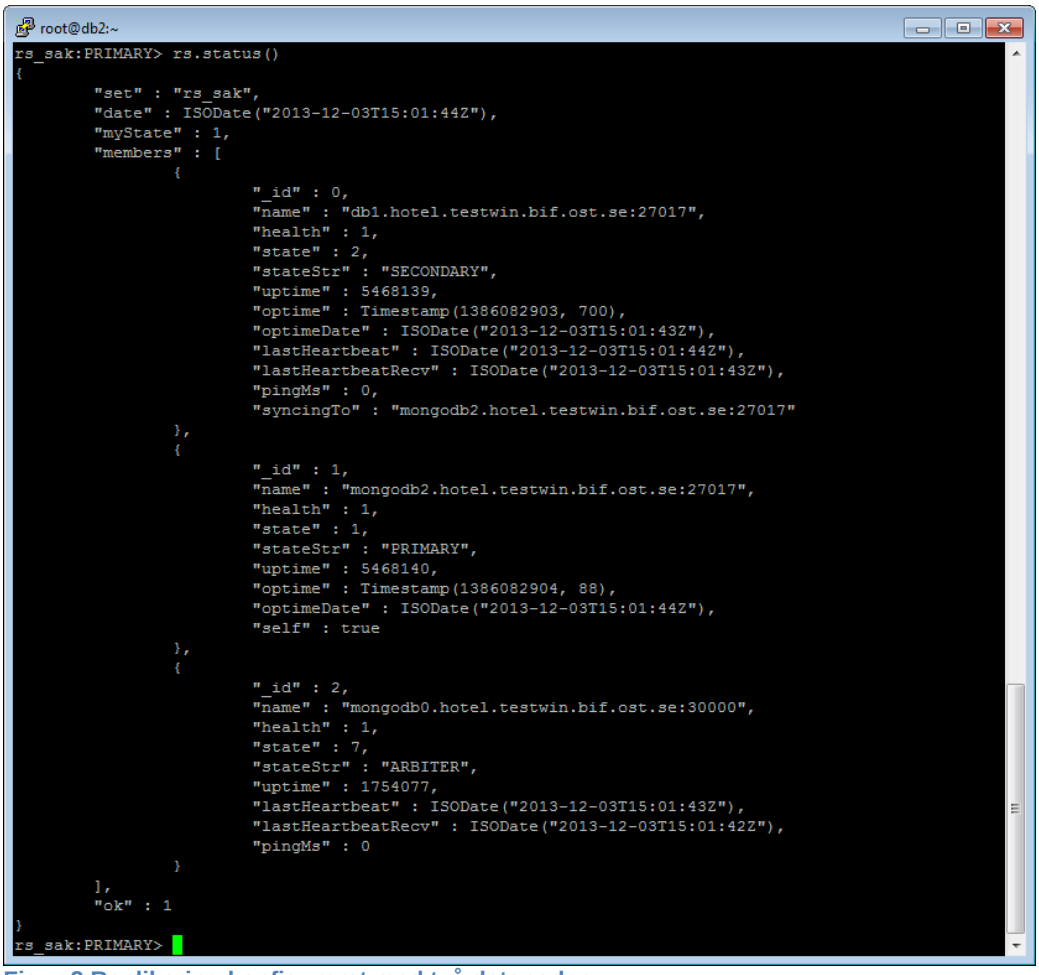
MongoDB är en så kallad dokumentorienterad databas eller NoSQL-databas. Den används inom tre funktioner i Säkerhetstjänster: Systemloggar, Loggrapportering och Arkivsökning. Med rekommenderad konfiguration och indexering blir denna lösning avsevärt mycket snabbare än tidigare lösning med Berkeley DB XML.

Med rekommenderad konfiguration menas en uppsättning med replikering av data över minst två noder, en *Primary*-nod och en *Secondary*-nod. Förutom det faktum att data replikeras för hög åtkomst så kommer Säkerhetstjänster att skriva data till den primära noden men begära att få läsa från en sekundär. Det innebär att lasten för skriv och läsoperationer fördelas på flera maskiner.
En övervakare, *Arbiter*, som håller reda på datanodernas hälsa installeras på en tredje nod.



Figur 1 Replikering i MongoDB

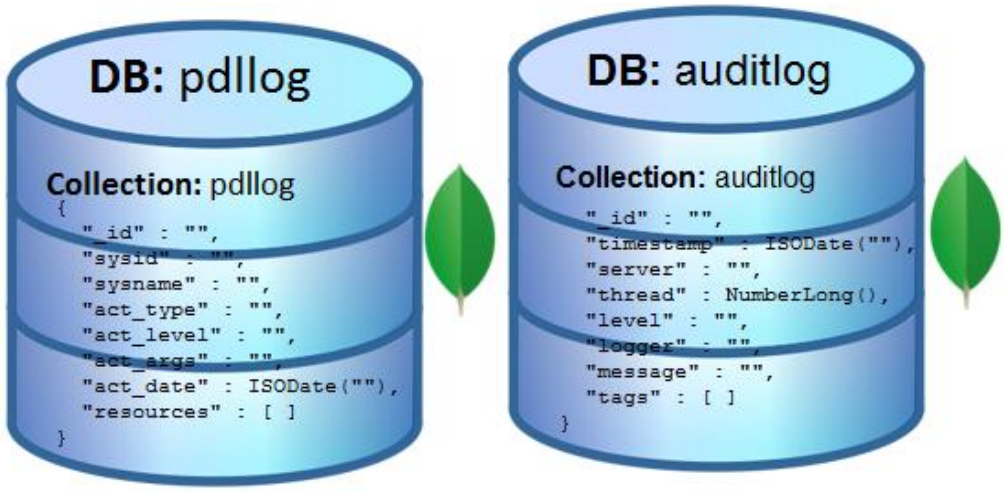
Installation av MongoDB och konfiguration av replikering och säkerhet görs enligt användarhandboken för Säkerhetstjänster. Förutom replikering måste databaser skapas, användare för Säkerhetstjänster definieras per databas och autentisering mellan noderna konfigureras.



Figur 2 Replikering konfigurerat med två datanoder

Indexering är också direkt avgörande för prestanda. För Systemloggar och Arkivsökning sker detta från Säkerhetstjänster men för Loggrapportering krävs att index skapas manuellt eller med hjälp av skriptfil som följer med installationen.

Var och en av de tre tjänsterna bör ha en egen instans av databasen. Installationen kommer att förvänta sig att tre databaser skapas, *auditlog* för Systemloggar, *pdllog* för Loggrapportering och *archivesearch* för Arkivsökning. Inom varje databas skapas sedan sedan en *”collection”* för varje konfiguration som är installerad.

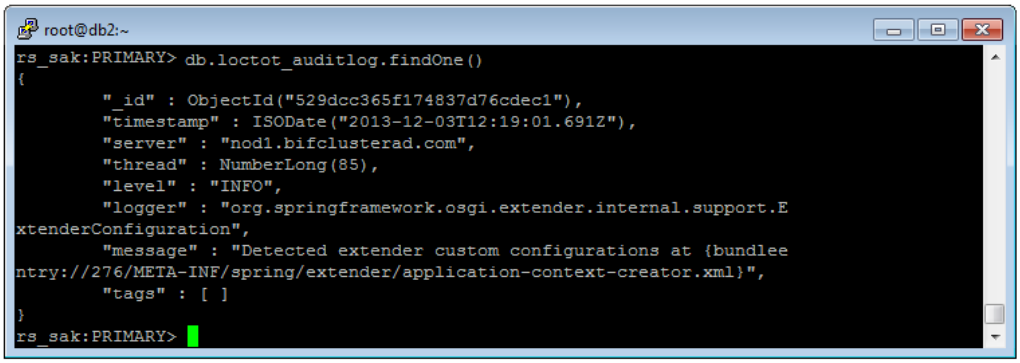


Figur 3 Databaser och Collections i MongoBD

Skapandet av databaser och användare som matchar konfigurationen av Säkerhetstjänster måste också göras manuellt eller med medföljande skript innan tjänsten startas.

- **System-loggar**. Alternativ källa till loggning av händelser i systemet. Detta görs i databasen ”*auditlog*”. Denna måste skapas innan systemet startas med en användare som matchar konfigurationen av modulen *com.logica.se.iac.logging.mongo*

  Installationen kommer endast att skapa databasen och en användare för Säkerhetstjänster. *Collection* och index kommer Säkerhetstjänster själv att skapa vid behov.



Figur 4 ”Schema” för databasen auditlog

- **PDL-loggar**. Databas med namnet ”*pdllog*” som håller data för Loggrapportering av verksamhetsloggar. Denna databas måste skapas innan systemet startas med en användare som matchar konfigurationen av modulen *com.logica.se.bif.logreport.store.online.mongodb.impl*

  Installation med hjälp av databasskriptet kommer att skapa databasen, användare för Säkerhetstjänster och en *collection* med korrekt indexering.

  Exempel på en loggpost som lästs in som dokument i collection pdllog:

```
{
   "_id" : "1558d6c1-a3b0-4979-ab32-98315cd6cf83",
   "sysid" : "systemId-1",
   "sysname" : "systemName",
   "act_type" : "Läsa",
   "act_level" : "activityLevel",
   "act_args" : "activityArgs",
   "act_date" : ISODate("2013-07-29T22:49:20.169Z"),
   "act_purpose" : "Vård och behandling",
   "uid" : "192109259321",
   "uname" : "Anders Andersson",
   "upid" : "194004884898",
   "assignment" : "assignment",
   "title" : "Läkaren",
   "cpid" : "Careprovider-4",
   "cpname" : "Careprovider-4",
   "cuid" : "CU-1215",
   "cuname" : "Careprovider-4- Careunit-1215",
   "resources" : [
      {
        "res_type" : "AutoSystemTest",
        "pid" : "192506244975",
        "pname" : "Sven Svensson",
        "cpid" : "Careprovider-4",
        "cpname" : "Careprovider-4",
        "cuid" : "CU-1215",
        "cuname" : "Careprovider-4-CU-1215"
       },
       {
         "res_type" : "AutoSystemTest",
         "pid" : "192506244975",
         "pname" : "Sven Svensson",
         "cpid" : "Careprovider-4",
         "cpname" : "Careprovider-4",
         "cuid" : "CU-1215",
         "cuname" : "Careprovider-4-CU-1215"
           }
   ]
}
```

Strukturen i detta MongoDB-dokument mappar mot schemat i loggarkivet enligt tabellen nedan.

Archive.xml	MongoDB PDL Log
<log:LogId>	_id
<log:SystemId>	sysid
<log:SystemName>	sysname
<log:ActivityType>	act_type
<log:ActivityLevel>	act_type
<log:ActivityArgs>	act_args
<log:StartDate>	act_date
<log:Purpose>	act_purpose
<log:UserId>	uid
<log:Name>	uname
<log:PersonId>	upid
<log:Assignment>	assignment
<log:Title>	title
<log:CareProviderId>	cpid
<log:CareProviderName>	cpname
<log:CareUnitId>	cuid
<log:CareUnitName>	cuname
<log:Resource>	resources
<log:ResourceType>	resources.res_type
<log:CareProviderId>	resources.cpid
<log:CareProviderName>	resources.cpname
<log:CareUnitId>	resources.cuid
<log:CareUnitName>	resources.cuname
<log:PatientId>	resources.pid
<log:PatientName>	resources.pname

Figur 5 Mappning XML-schema mot MongoDB-dokumentstruktur för PDL-loggar

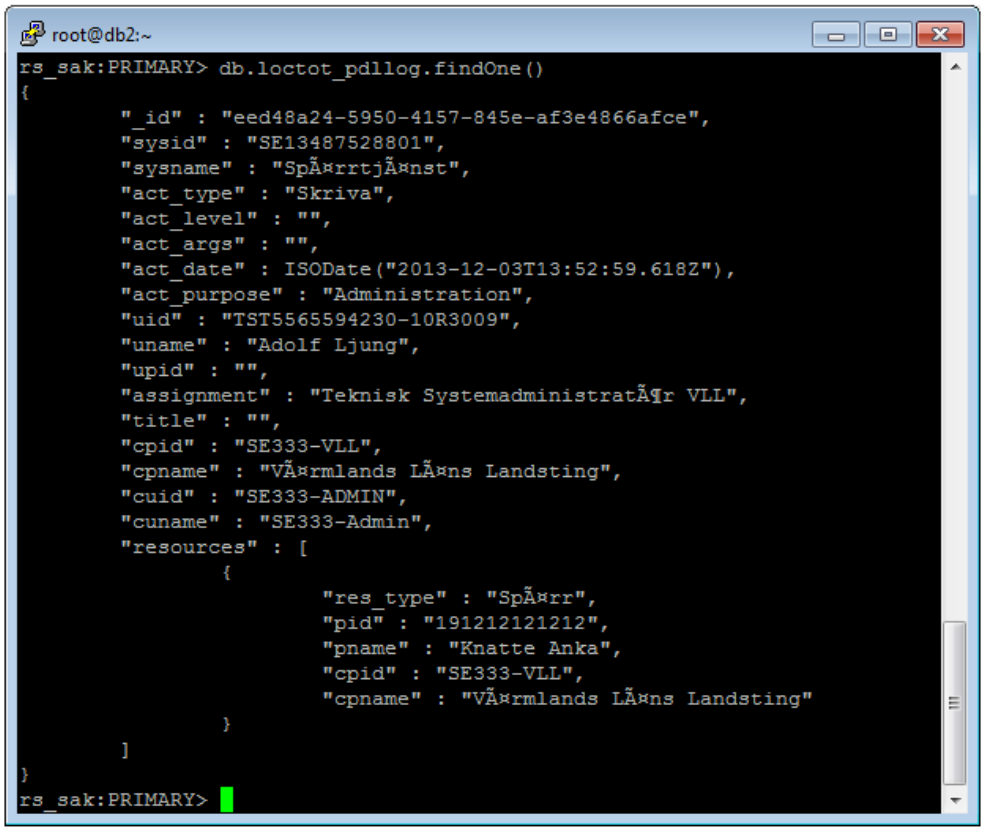
De index som installationsskriptet lägger på *pldlog-collection* är viktiga för prestandan vid uttag av loggrapporter. De baseras på de sökningar i MongoDB som görs för de olika rapporttyperna och är av sorten sammansatta index (*compund indexes*).

Index läggs in per *collection* med funktionen ensureIndex så dessa kommandon kommer att exekveras av installationsskriptet:

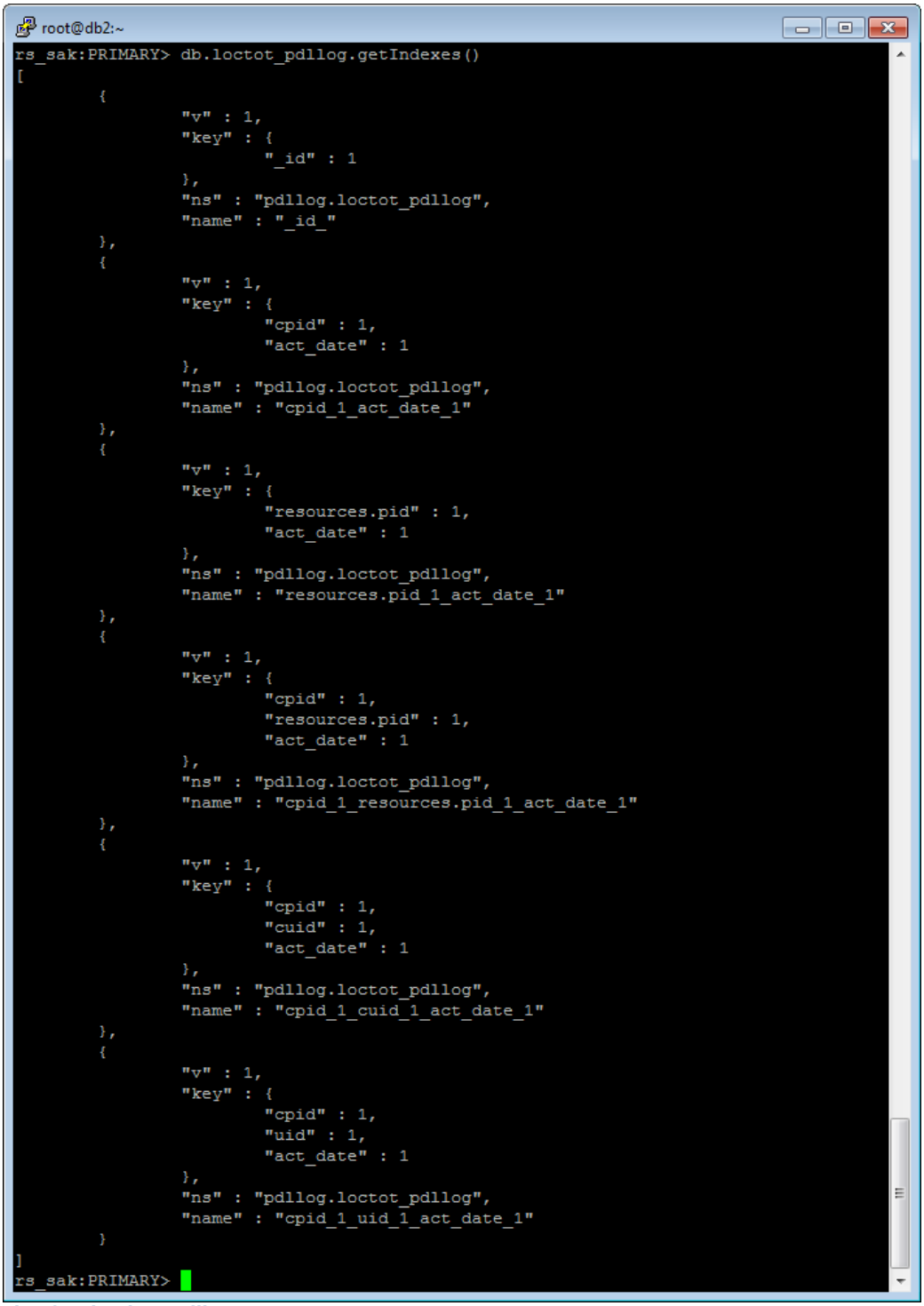
```
db.log.ensureIndex({"cpid": 1, "act_date" : 1})
db.log.ensureIndex({"resources.pid": 1, "act_date" : 1})
db.log.ensureIndex({"cpid": 1, "resources.pid":1, "act_date":1})
db.log.ensureIndex({"cpid": 1, "cuid": 1, "act_date" : 1})
db.log.ensureIndex({"cpid": 1, "uid": 1, "act_date" : 1})
```

Index	Rapport	Info
{"cpid": 1, "resources.pid": 1, "act_date" : 1}	Patient	Åtgärder avseende viss patient (inom egen vårdgivare)
{"cpid": 1, "uid": 1, "act_date" : 1}	Personal	Åtgärder som viss personal har vidtagit (inom egen vårdgivare)
{"cpid": 1, "act_date" : 1}	Vårdgivare	Åtgärder rörande all personal inom egen vårdgivare
{"cpid": 1, "cuid": 1, "act_date" : 1}	Patient, vårdenhet	Åtgärder avseende viss patient utifrån angiven vårdenhet (inom egen vårdgivare)
{"resources.pid": 1, "act_date" : 1}	Patientperspektiv	Lista för angiven patient, vilka vårdgivare som har haft åtkomst till information

Figur 6 Mappning mellan rapporttyp och index för *collection* i databasen PDLLOG



Figur 7 "Schema" för databas pdllog



Figur 8 Index för databas pdllog

- **Arkivsökning**. Detta är en databas med namnet ”*archivesearch*” som används vid sökning i arkiverade loggarkiv. Arkivsökning som funktion resulterar egentligen i en XML-fil på samma format som en PDL-loggrapport och därför kommer schema och index för *collections* i denna databas att vara identiska med de för databasen *pdllog*. Den stora skillnaden är att för arkivsökningen hämtas data från komprimerade arkiv och funnen data lagras i databasen bara temporärt. Från dessa data skapas sedan en rapport men när den är skapad tas *collection* bort.

  Databasen måste dock skapas, med en användare som matchar konfigurationen av modulen *com.logica.se.bif.logreport.archive.store.mongodb.impl*, innan systemet startas. Installationen kommer endast att skapa databasen och en användare för Säkerhetstjänster. *Collection* och index kommer Säkerhetstjänster själv att skapa vid behov.