

AIS data from Kystverket

-

Lambda architecture with Spark Streaming and StreamSets

Demo to show:

- how streaming data can be used on Spark/Hadoop in a lambda architecture («real time» + batch)
- how an ETL pipeline can be built with StreamSets
- how Solr and Hue can be used to get a quick view of large batch data

Data prepp
(not always so fun)



Analytics
(fun...)

Data source:


KYSTVERKET

[Hjem](#) [Maritime tjenester](#) [Beredskap](#) [Maritim infrastruktur](#) [Regelverk](#) [Kystkultur](#) [Om Kystverket](#) [Presse](#)

Forside

[GÅ TIL NYHETSARKIVET](#)

Publisert: 20.09.2016 10:37:32, Oppdatert: 20.09.2016 14:55:49 - Anne Grethe Nilsen



Åpner AIS for publikum

Nå kan publikum følge skipstrafikken live via Kystinfo, som nylig ble koblet opp mot Kystverkets AIS nettverk, AIS Norge. Nettverket viser skipsbevegelser i sanntid og har tidligere kun vært tilgjengelig for offentlige myndigheter og havner.

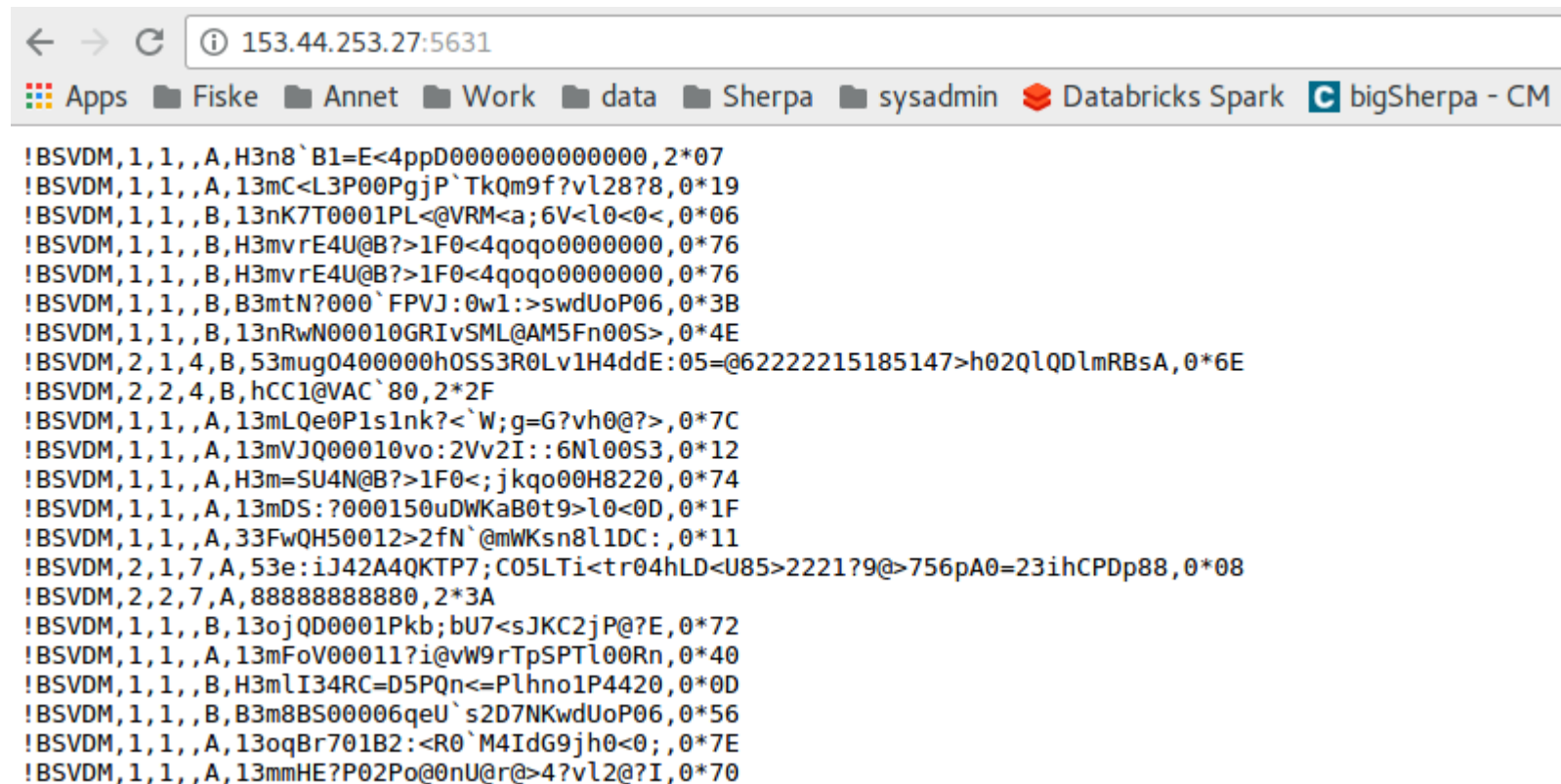
– AIS-data er verdifulle data med stor samfunnsnyttig verdi. Ved å gjøre disse dataene tilgjengelig for publikum håper vi å bidra til verdiskapning, effektivisering og innovasjon både i privat og offentlig sektor, sier sjøsikkerhetsdirektør Arve Dimmen i Kystverket.

Se skipsbevegelser i sanntid

AIS viser skipsbevegelser til skip over 45 meter opptil 12 nautiske mil fra kysten. AIS-data hentes inn fra Kystverkets 50 basestasjoner langs kysten som fanger opp signaler fra alle skip over 300 bruttotonn som seiler i internasjonal fart.

Data source 1:

- Address: 153.44.253.27 : 5631



```
!BSVDM,1,1,,A,H3n8`B1=E<4ppD0000000000000,2*07
!BSVDM,1,1,,A,13mC<L3P00PgjP`TkQm9f?v128?8,0*19
!BSVDM,1,1,,B,13nK7T0001PL<@VRM<a;6V<l0<0<,0*06
!BSVDM,1,1,,B,H3mvrE4U@B?>1F0<4qoqo0000000,0*76
!BSVDM,1,1,,B,H3mvrE4U@B?>1F0<4qoqo0000000,0*76
!BSVDM,1,1,,B,B3mtN?000`FPVJ:0w1:>swdUoP06,0*3B
!BSVDM,1,1,,B,13nRwN00010GRIVSML@AM5Fn00S>,0*4E
!BSVDM,2,1,4,B,53mug0400000h0SS3R0Lv1H4ddE:05=@62222215185147>h02QlQDlmRBsA,0*6E
!BSVDM,2,2,4,B,hCC1@VAC`80,2*2F
!BSVDM,1,1,,A,13mLQe0P1s1nk?<`W;g=G?vh0@?>,0*7C
!BSVDM,1,1,,A,13mVJQ00010vo:2Vv2I::6Nl00S3,0*12
!BSVDM,1,1,,A,H3m=SU4N@B?>1F0<;jkqo00H8220,0*74
!BSVDM,1,1,,A,13mDS:?000150uDWKaB0t9>l0<0D,0*1F
!BSVDM,1,1,,A,33FwQH50012>2fN`@mWKsn8l1DC:,0*11
!BSVDM,2,1,7,A,53e:iJ42A4QKTP7;C05LTi<tr04hLD<U85>2221?9@>756pA0=23ihCPDp88,0*08
!BSVDM,2,2,7,A,88888888880,2*3A
!BSVDM,1,1,,B,13ojQD0001Pkb;bU7<sJJC2jP@?E,0*72
!BSVDM,1,1,,A,13mFoV00011?i@vW9rTpSPTl00Rn,0*40
!BSVDM,1,1,,B,H3mLI34RC=D5PQn<=Plhno1P4420,0*0D
!BSVDM,1,1,,B,B3m8BS00006qeU`s2D7NKwdUoP06,0*56
!BSVDM,1,1,,A,13oqBr701B2:<R0`M4IdG9jh0<0;,0*7E
!BSVDM,1,1,,A,13mmHE?P02Po@0nU@r@>4?v12@?I,0*70
```

- NMEA – decoded, converters available in multiple languages

Data source 1 decoding:

```
In [1]: import ais

In [2]: ais.decode("13mLQe0P1s1nk?<`W;g=G?vh0@?>",0)
Out[2]:
{'cog': 342.0,
 'id': 1L,
 'mmsi': 257368500L,
 'nav_status': 0L,
 'position_accuracy': 0L,
 'raim': False,
 'repeat_indicator': 0L,
 'rot': -731.386474609375,
 'rot_over_range': True,
 'slot_number': 974L,
 'slot_timeout': 4L,
 'sog': 12.300000190734863,
 'spare': 0L,
 'special_manoeuvre': 0L,
 'sync_state': 0L,
 'timestamp': 24L,
 'true_heading': 511L,
 'x': 25.952383041381836,
 'y': 70.97503662109375}

In [3]: █
```

Useful parameters:

- 'mmsi': ID
- 'x': longitude
- 'y': latitude
- 'cog': course over ground
- 'sog': speed over ground
- 'rot': rate of turn

Timestamp??

Ship metadata??

Data source 2:

- HBase table with ship metadata

```
hbase(main):008:0> get 'mmsiShipInfo','259402000'
COLUMN                                CELL
P:imo                                timestamp=1478526009855, value=9144093
P:mmsi                                timestamp=1478526009855, value=259402000
P:name                                timestamp=1478526009855, value=BASTO 2
P:type                                timestamp=1478526009855, value=Ro-Ro/Passenger Ship
4 row(s) in 0.0330 seconds
```

Data flow:

bigSherpa – Hadoop Cluster

