SPIRE

9 SAP HANA ANALYSIS

By Dean Keithly

The purpose of this section is to perform an analysis determining how HANA could prove their methods are faster.

The core problem is a need to calculate whether any two spacecraft will collide. Generally speaking, a spacecraft maneuver must be performed when the probability of two spacecraft colliding reaches above some pre-defined threshold, a percentage. I postulate that this collision threshold is achieved when the position uncertainty covariance surfaces at that threshold of the two different spacecraft intersect.

A spacecraft position has a temporally varying covariance matrix dependent upon its orbital parameters and position. This can be loosely described by x,dx1,dx2,y,dy1,dy2,z,dz1,dz2,xd,dxd1,dxd2,yd,dyd1,dyd2,zd,dzd1,dzd2,t. If each of these parameters are long doubles, equivalent to 8 bytes, then each spacecraft position has 19 parameters (keplerian elements -¿ 6 x 3 + 1 = 19) and 152 bytes. This could be reduced if bi-directional components dyd1 and dyd2 are (or are assumed to be) identical.

Let us take the fastest moving spacecraft in orbit to determine the number of point-positions we will need. This is a low-earth orbit spacecraft. I will assume the time between orbit characterization revisits is 1 month. We can approximate the number of orbits in a month as

$$1mo * \frac{30d}{1mo}\frac{24hr}{1d}\frac{60min}{1hr}\frac{1orbit}{90min} = 480orbits. \tag{39}$$

If $R_\oplus = 6,356$km, then by $2\pi R_\oplus$ is 39,935,892.08 m per orbit. The total distance traversed by a spacecraft in a month is therefore 19,169,228,198.4 m.

Lets just say the collision radius is 20m. If so, we can calculate the number of spacecraft positions that must be calculated. We get 19,169,228,198.4m/20m = 958,461,409.92 points (approximate as 10^9 points). These are the number of location data points for 1 spacecraft. Space-track says there are approximately 134,114,651 ($\approx 1.5 \times 10^5$) TLEs in space and 21,000 large object with 17266 of them being spacecraft.

If we cross-compare every space object (SC), SC1 and SC2 have $10^9 \times 10^9$ positions to compare. SC1 to all other SCi for all $10^9 \times 10^9 \times (1.5 \times 10^5)^2/2 = 7.5 \times 10^{27}$.

The total number of bytes in storage are $152 \times 7.5 \times 10^{27} = 1.14 \times 10^{30}$ (Note: 1 petabyte is 2^{50} bytes)

Storing state vectors for each object is therefore infeasible for SAP HANA since the largest volume of data they currently track is 12.67 petabytes (12.67×10^{15} bytes or 1.267×10^{16}).

9.1 Efficiency of a Gridded System

GEOsynchronous orbit has an altitude of 35,786 km. Total volume is $4.582890417 \times 10^{22}$ if each bin is a 1m cube.

The spacecraft state-vector positions are then transformed into probabilities of being in each cube at each time dt. Data is now saved as (bin index, space object #, time, (maybe probability)). To sort for collisions, this is as simple as

```
Result: Write here the result
initialization;
for bin do
    for time t do
        S = SC in bin at t;
        if Π_{i∈S}(p_i) > thresh then
            Collision;
        else
            continue;
        end
    end
end
```

Algorithm 1: Gridded System Collision Frequency

The schema for this is to give each bin a unique ID number. The id could be devolved into 3 numbers, of $\pm 35,786,000$m uniquely identifying each 1m bin. To determine whether collisions occur, look at each bin, look for the set of all spacecraft entering that bin, if the times they enter that bin are within threshold, then a collision occurs. The number of entries in the database would be, at maximum, $19,169,228,000 \times 134,114,651$ (the maximum number of cubes the spacecraft go through, the maximum number of objects tracked), totalling $2.570874323 \times 10^{18}$ (this could increase by as much as 2x based on the minimum and maximum number of cubes a vector passing through a 2x2x2 cube bank can intersect). $2.570874323 \times 10^{18}$ is still infeasible since the current SAP HANA limitation is 10^{16} bytes of storage.

If we relaxed the cube size constraints from 1m cube to 10m cube, we reduce $19,169,228,000$ by 3 orders of magnitude resulting in a final number of entries in the database of $2.570874323 \times 10^{15}$ data elements. A data element would be (xbin,ybin,zbin,SCid,time) with types (float, float, float, float, float) and sizes (4 bytes, 4 bytes, 4 bytes, 4 bytes, 4 bytes) totalling 20 bytes per element with minor possible savings. We now know we need $5.141748646 \times 10^{16}$ bytes to store all the required spacecraft position data.

This fits within the current maximum capability of SAP HANA as indicated by the quoted 1.267×10^{16}.

9.2 Assumption Validity: Bin storage cubes vs spheres

Additional memory simplification is possible. By cutting off the cube regions beyond Geosynchronous orbit, it is possible to save $1 - \frac{\frac{4}{3}\pi 35,786,000^3}{(2\times 35,786,000)^3} = 1 - \frac{\pi}{6} =$

0.4764 (could save 47.64%). Additional cube storage reduction can occur by removing MEO cube locations since the spacecraft density here is low. The Earth can also be subtracted out but this represents a small gain. Additionally, these calculations assume tracking of all objects but if we could ignore small objects, we could substantially reduce the number of tracked objects. The number of cubes spacecraft move through in GEO is much smaller so it is possible to use different time scales for collision detection.

9.3 Assumption Validity: 1 month of orbit

From https://celestrak.com/columns/v04n05/: "How often are element sets generated? New element sets are generated by NORAD on an as-needed basis rather than according to an established timetable. How often these updates occur depends upon a number of factors such as the orbit type or maneuvering capability of the satellite. For example, a satellite in low-earth orbit—such as the US space shuttle—would have its element sets updated several times a day because of the somewhat unpredictable results of atmospheric drag as it varies its attitude and the maneuvering being performed. A satellite in a low-drag orbit which doesn't maneuver—such as LAGEOS II—might only need updates once or twice a week. Objects such as rocket bodies, defunct payloads, or other space debris, won't be updated as frequently, either—unless there is a prediction of a close approach with an operational payload. Special-interest objects—such as a large object reentering the earth's atmosphere—normally get special treatment."
This indicates my assumed 30d from TLEs is at likely 4 times as long as it should be.

9.4 Potential enhancement: Clustering Debris

Debris from space object collisions is often clustered in the same oribt.

9.5 Full Data Pipeline

Ground observations $\Rightarrow$ orbit fitting models $\Rightarrow$ TLE $\Rightarrow$ detailed orbit propegator $\Rightarrow$ spacecraft state vectors $\Rightarrow$ state vectors to position bins in SAP HANA $\Rightarrow$ identify collisions