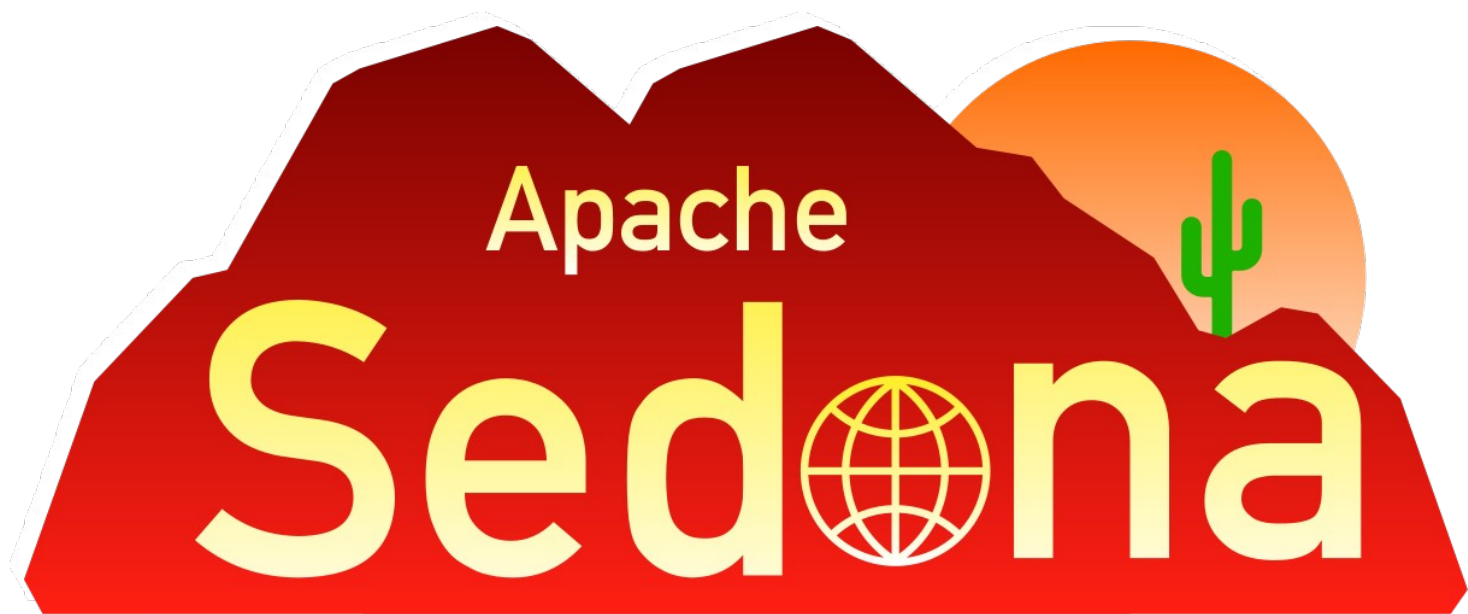


Apache Sedona

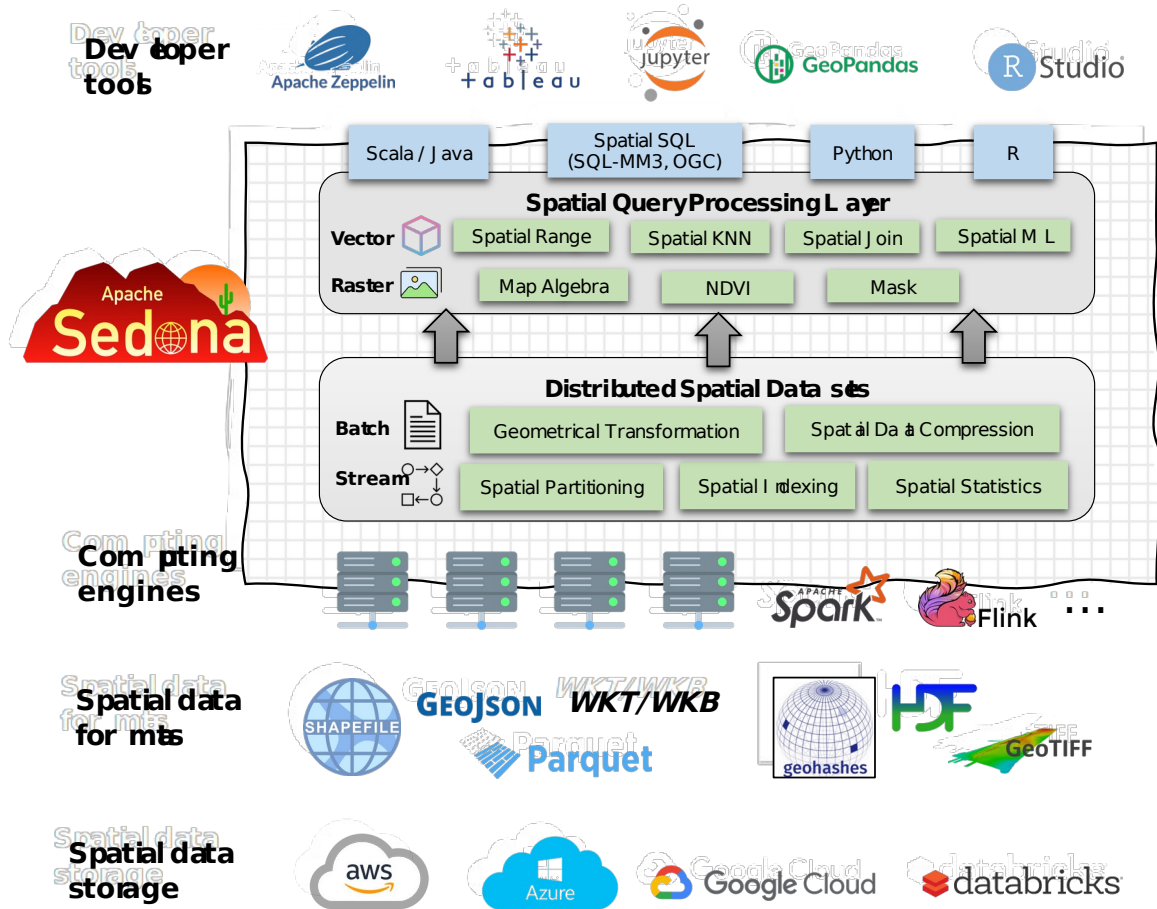
Sistema de computação em cluster para processamento em larga
escala de dados espaciais



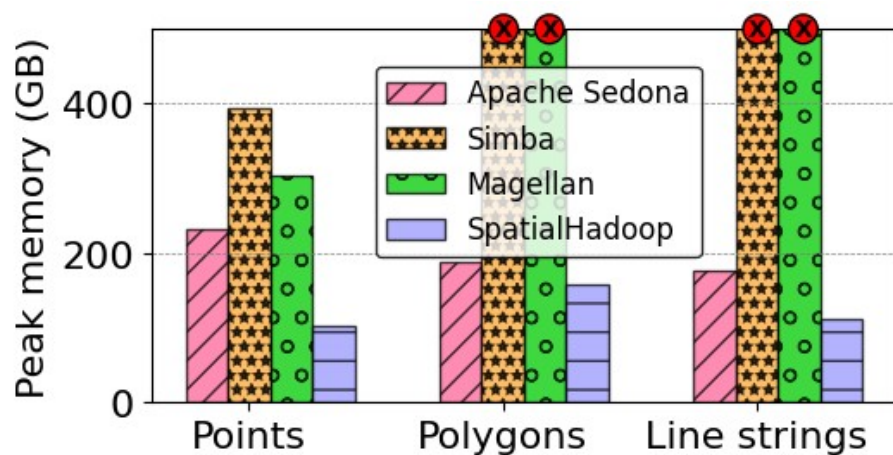
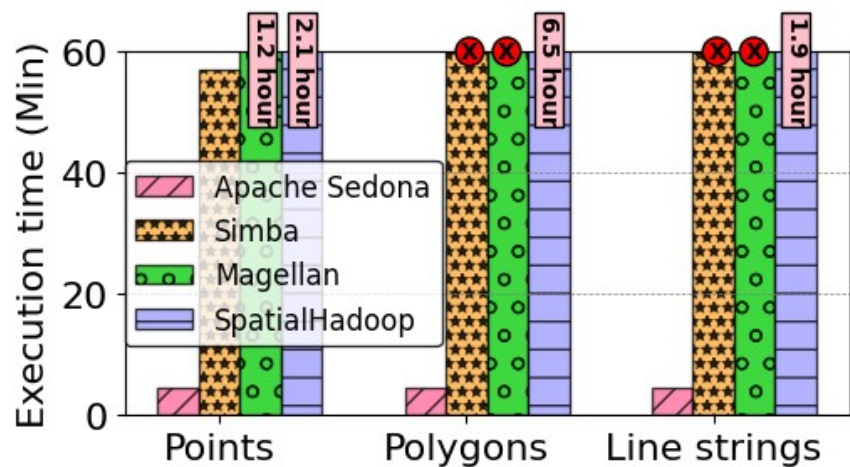
Apache Sedona

Apache Sedona é um sistema de computação em cluster para processamento de dados espaciais em larga escala. Sedona estende os sistemas de computação em cluster existentes, assim como Apache Spark e Apache Flink, com um conjunto de datasets espaciais distribuídos prontos para uso e Spatial SQL que carregam, processam e analisam com eficiência dados espaciais em larga escala nas máquinas.

Arquitetura



Vantagens



Análise Espacial

Apache Sedona permite integração com as linguagens Java/Scala, Python e R. Além do suporte a Spatial SQL com o GeoSpark, permitindo assim, desenvolvimento de apps para análise espacial e mineração de dados em ambiente de nuvem.



Exemplo de uso

O ambiente de teste foi desenvolvido localmente com a linguagem Scala na versão 3.2.1 com engine Spark e uso de formatos WKT para representação de objetos geométrico vetoriais.

- Fonte: 50 pontos turísticos mais visitados do mundo e fronteiras de países
- Dependências:

```
// sedona
libraryDependencies += "org.apache.sedona" % "sedona-python-adapter-3.0_2.13" % "1.3.1-incubating",
libraryDependencies += "org.apache.sedona" % "sedona-viz-3.0_2.13" % "1.3.1-incubating",
libraryDependencies += "org.datasyslab" % "geotools-wrapper" % "1.3.0-27.2",
// spark
libraryDependencies += "org.apache.spark" % "spark-core_2.13" % "3.3.1",
libraryDependencies += "org.apache.spark" % "spark-sql_2.13" % "3.3.1",
```

Consultas

```
SELECT name,x,y,SQRT(  
    POW(abs(x-(-47.9341)),2)+POW(abs(y-(-15.7792)),2)  
) as distance  
FROM famous_points  
ORDER BY distance ASC  
LIMIT 5
```

```
SELECT name,x,y,ST_Distance(  
    ST_GeomFromWKT('POINT (-47.9341 -15.7792)'), st_point(x,y)  
) AS distance  
FROM famous_points  
ORDER BY distance ASC  
LIMIT 5
```

name	x	y	distance
Disney's Hollywood Studios - Florida	-81.5612862	28.3580628	55.4877069196929
Epcot - Florida	-81.5515921	28.3764687	55.4964760452204
Disney's Animal Kingdom - Florida	-81.6057808	28.356256	55.51324674548431
Islands of Adventure - Universal Orlando	-81.470008	28.4743207	55.52505038025443
Universal Studios - Florida	-81.4841515	28.4754235	55.53447268840684

Pontos turísticos nos E.U.A.

Seleciona todos os pontos turísticos que estão nos Estados Unidos, onde as fronteiras do país é formado por multipoligonos em formato WKT.

```
SELECT name,x,y
FROM famous_points
WHERE ST_Contains([
    ST_GeomFromWKT((
        SELECT wkt
        FROM countries
        WHERE countries.country == 'United States of America'
    )),
    st_point(x,y)
])
```


name	x	y
Bourbon Street - New Orleans	-90.069155	29.9547116
Metropolitan Museum of Art - New York City	-73.9654327	40.7794366
Universal Studios - Florida	-81.4841515	28.4754235
Lake Mead - Nevada	-115.469129	36.1435231
Smithsonian National Air and Space Museum - Washington D.C.	-77.0220566	38.8881601
Disney's California Adventure	-117.9230477	33.806112
Islands of Adventure - Universal Orlando	-81.470008	28.4743207
Smithsonian National Museum of Natural History - Washington D.C.	-77.0282541	38.8912662
Pier 39 - San Francisco	-122.4139409	37.808673
Navy Pier - Chicago	-87.6072831	41.8918633
Great Smoky Mountains - National Park - Tennessee	-84.6078412	35.6130503
Disney's Hollywood Studios - Florida	-81.5612862	28.3580628
Disney's Animal Kingdom - Florida	-81.6057808	28.356256
Epcot - Florida	-81.5515921	28.3764687
Golden Gate Park - San Francisco	-122.4884025	37.7694208
Disneyland Park - Anaheim	-117.9211629	33.8120918
Disneyworld's Magic Kingdom - Orlando	-81.5833949	28.418579
Faneuil Hall Marketplace - Boston	-71.0573032	42.360189
Grand Central Terminal - New York City	-73.9794181	40.7527262
Union Station - Washington D.C.	-77.0222416	38.8977859
Central Park - New York City	-73.9820024	40.7735718
Times Square - New York City	-74.0514606	40.7579747
Las Vegas Strip	-115.2014624	36.1174891



FIM