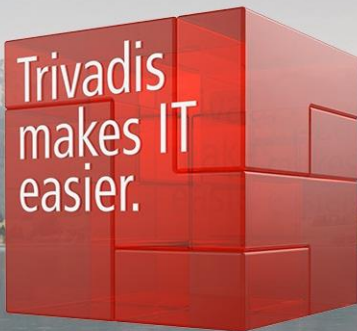


Approximate Query Processing

Christian Antognini



 @ChrisAntognini  antognini.ch/blog

BASEL ▪ BERN ▪ BRUGG ▪ DÜSSELDORF ▪ FRANKFURT A.M. ▪ FREIBURG I.BR. ▪ GENEVA
HAMBURG ▪ COPENHAGEN ▪ LAUSANNE ▪ MUNICH ▪ STUTTGART ▪ VIENNA ▪ ZURICH

trivadis
makes IT easier. ■ ■ ■

■ @ChrisAntognini

Senior principal consultant, trainer and partner at Trivadis

■ christian.antognini@trivadis.com

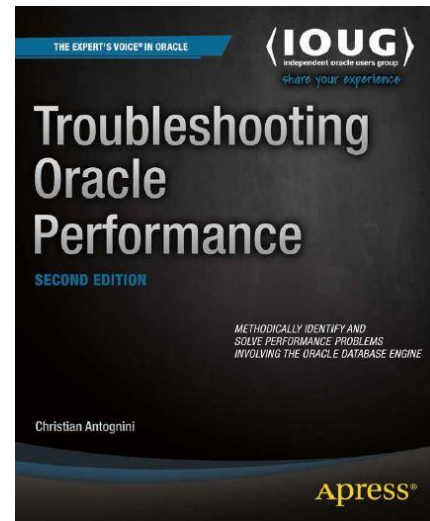
■ http://antognini.ch

Focus: get the most out of Oracle Database

- Logical and physical database design
- Query optimizer
- Application performance management

Author of Troubleshooting Oracle Performance (Apress, 2008/14)

OakTable Network, Oracle ACE Director



■ Agenda

1. 12.1.0.2 Features

2. 12.2 Features

12.1.0.2 Features

■ APPROX_COUNT_DISTINCT Function

Computing the number of distinct values can be resource intensive

In some situation performance is more important than precision

Many algorithms that *estimate* the number of distinct values have been developed

Oracle implemented one of this algorithms (HyperLogLog)

In case an estimated number of distinct values is acceptable, you can replace `COUNT(DISTINCT expr)` with `APPROX_COUNT_DISTINCT(expr)`

■ Test Case – Setup

100'000'000 rows stored in a 12.1.0.2 database

■ 12.7 GB

■ 24 columns

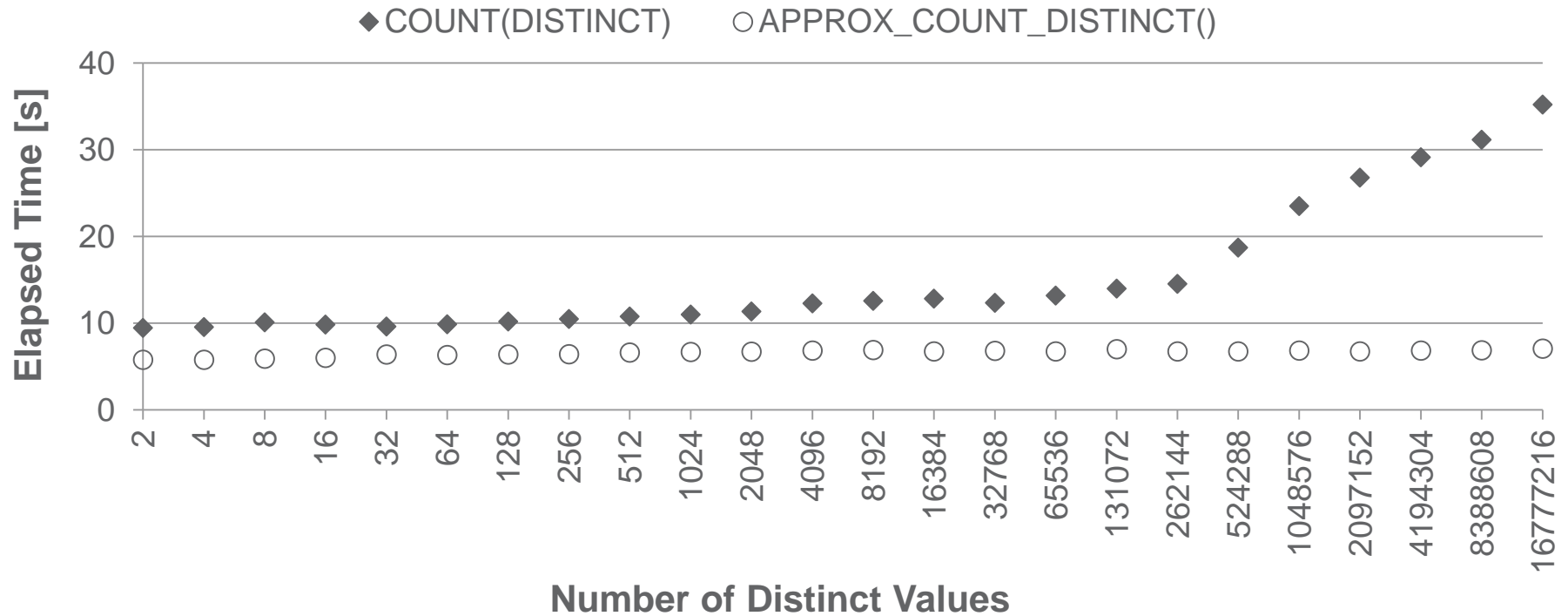
■ Number of distinct values is 2^m , where $\{m \in \mathbb{N}^* | m < 25\}$

Test queries:

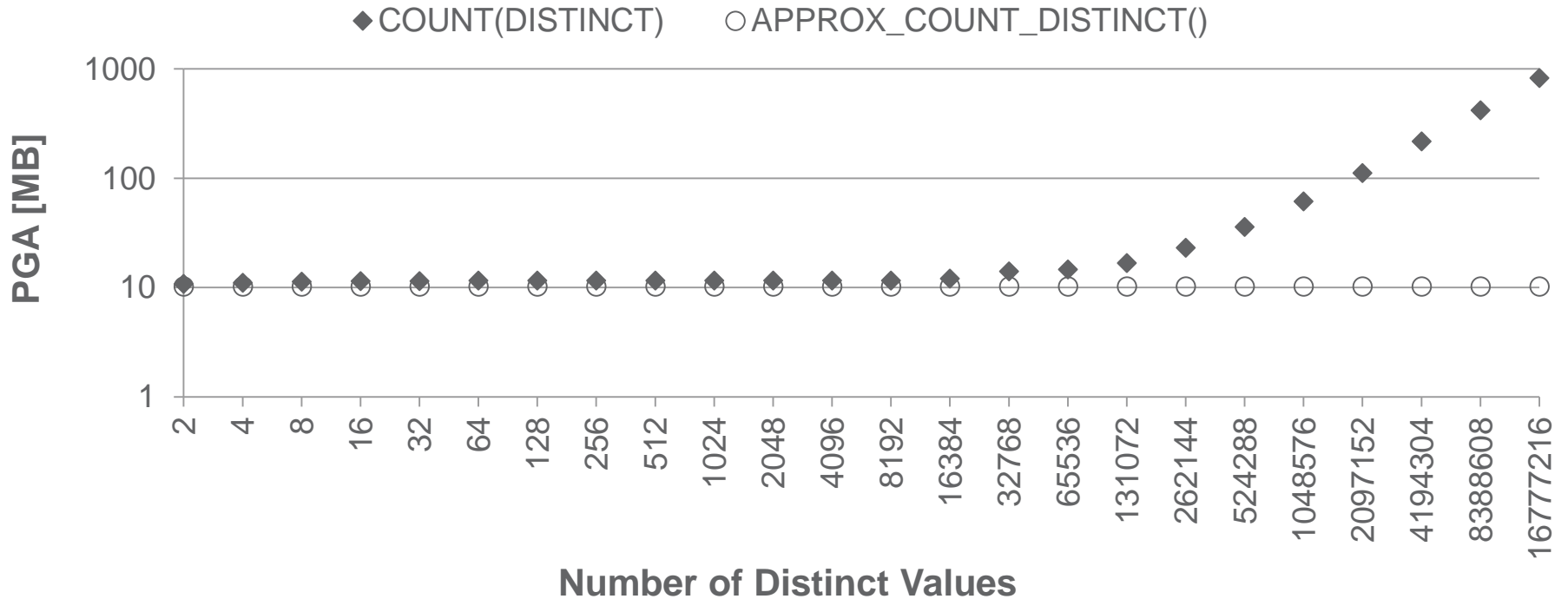
```
SELECT count(DISTINCT n_m) FROM t
```

```
SELECT approx_count_distinct(n_m) FROM t
```

Test Case – Elapsed Time



Test Case – PGA Utilization



■ Test Case – Precision



12.2 Features

■ New Features in 12.2

Base functions

- APPROX_MEDIAN
- APPROX_PERCENTILE

Approximate aggregate transformation and related parameters

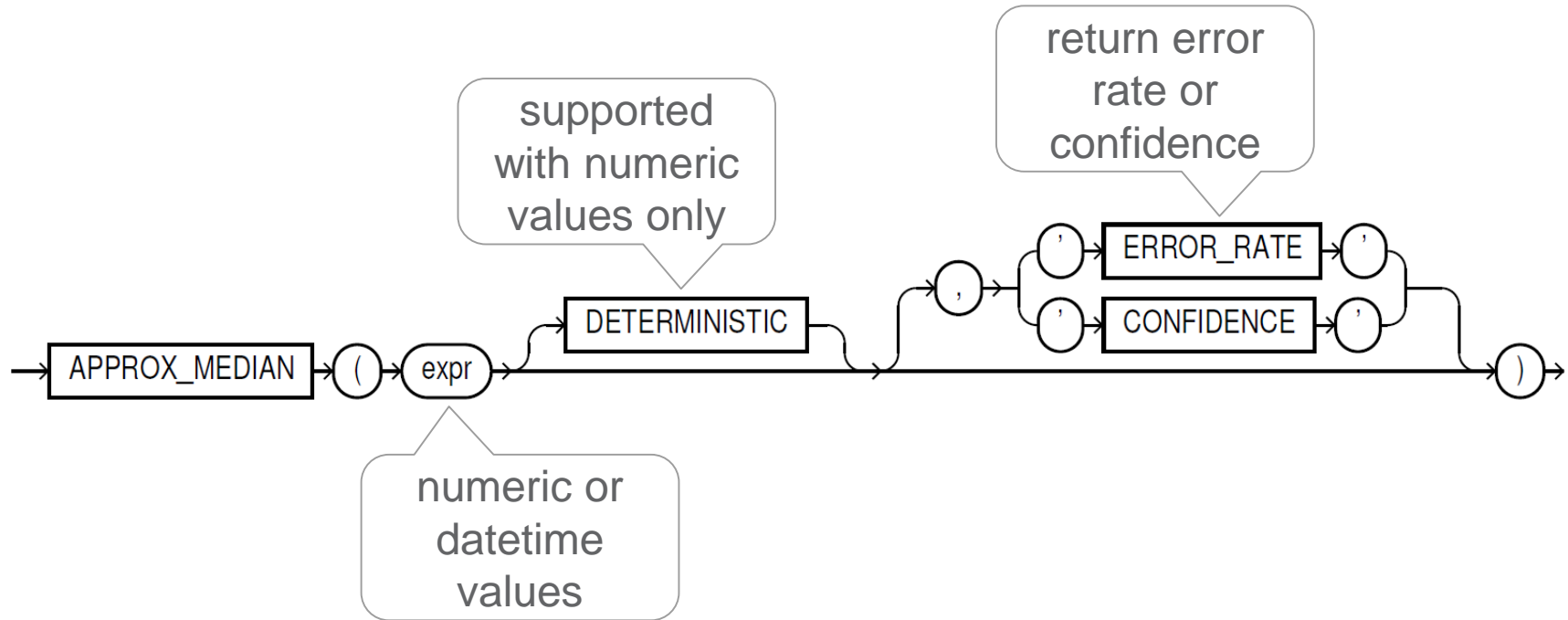
Support in materialized views

- Including query rewrite

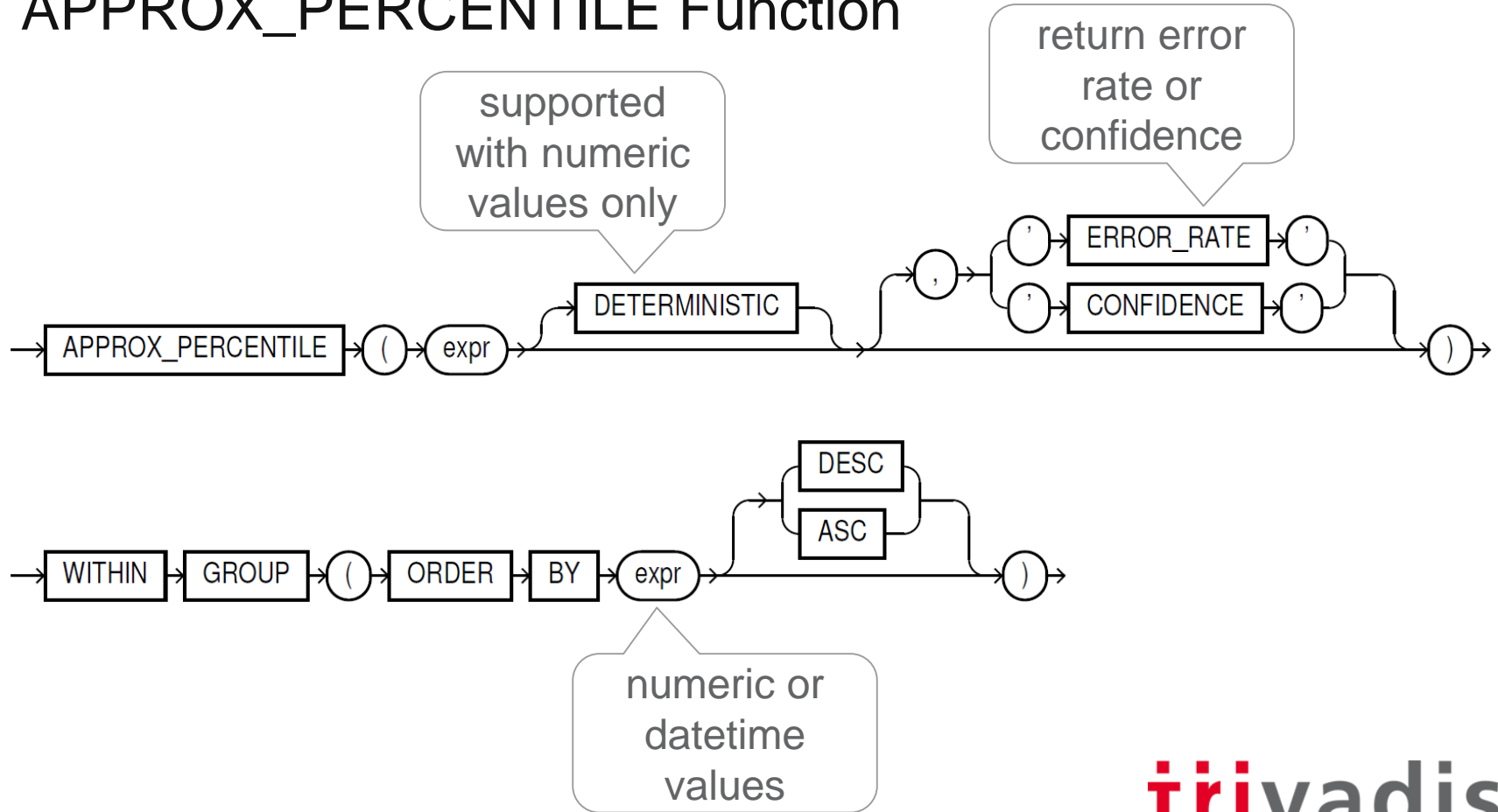
Helper functions

- APPROX_COUNT_DISTINCT_DETAIL
- APPROX_COUNT_DISTINCT_AGG
- TO_APPROX_COUNT_DISTINCT
- APPROX_PERCENTILE_DETAIL
- APPROX_PERCENTILE_AGG
- TO_APPROX_PERCENTILE

■ APPROX_MEDIAN Function



■ APPROX_PERCENTILE Function



■ Approximate Aggregate Transformation – Count Distinct

If `APPROX_FOR_COUNT_DISTINCT` is set to `TRUE` (default is `FALSE`), the optimizer applies the following transformation:

`COUNT(DISTINCT <expr>)` → `APPROX_COUNT_DISTINCT(<expr>)`

■ Approximate Aggregate Transformation – Percentile (1)

If APPROX_FOR_PERCENTILE is *not* set to the default value (NONE), the optimizer applies the following transformation:

MEDIAN(<expr>)

→ APPROX_PERCENTILE(0.5) WITHIN GROUP (ORDER BY <expr>)

■ Approximate Aggregate Transformation – Percentile (2)

If APPROX_FOR_PERCENTILE is set to PERCENTILE_CONT or ALL (default is NONE), the optimizer applies the following transformation:

```
PERCENTILE_CONT(<expr>) WITHIN GROUP (ORDER BY <expr>)  
→ APPROX_PERCENTILE(<expr>) WITHIN GROUP (ORDER BY <expr>)
```


■ Approximate Aggregate Transformation – Percentile (3)

If APPROX_FOR_PERCENTILE is set to PERCENTILE_DISC or ALL (default is NONE), the optimizer applies the following transformation:

```
PERCENTILE_DISC(<expr>) WITHIN GROUP (ORDER BY <expr>)  
→ APPROX_PERCENTILE(<expr>) WITHIN GROUP (ORDER BY <expr>)
```

■ Approximate Aggregate Transformation – Notes

The optimizer isn't able to apply the query transformation in case the OVER clause is specified

If APPROX_FOR_AGGREGATION is set to TRUE (default is FALSE) and provided the other parameters aren't set, all query transformations are enabled

■ Helper Functions and Materialized Views

The result of APPROX_COUNT_DISTINCT, APPROX_MEDIAN and APPROX_PERCENTILE can be stored in tables and materialized views

However, further aggregations are not possible!

The helper function has to be used to store approximations that has to be post-processed

- APPROX_*_DETAIL generates data that allows post-processing
- APPROX_*_AGG aggregates data generated by APPROX_*_DETAIL
- TO_APPROX_* converts in readable form the output of APPROX_*_AGG

■ Core Messages



- Useful features
- Sound implementation

Questions and Answers

Christian Antognini
Senior Principal Consultant

christian.antognini@trivadis.com

[@ChrisAntognini](https://twitter.com/ChrisAntognini)

