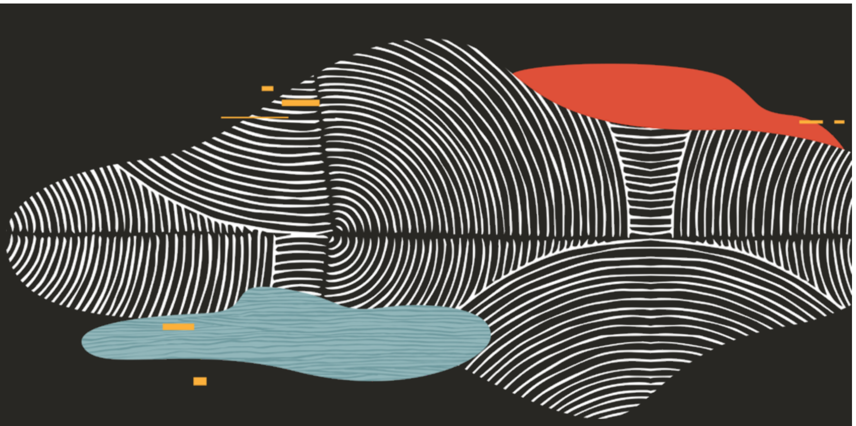
ORACLE

Protecting Critical OLTP Workloads in a Mixed Workload Environment

Mihajlo Tekic

John Zimmerman

Real-World Performance Team
Oracle Database Development



Agenda

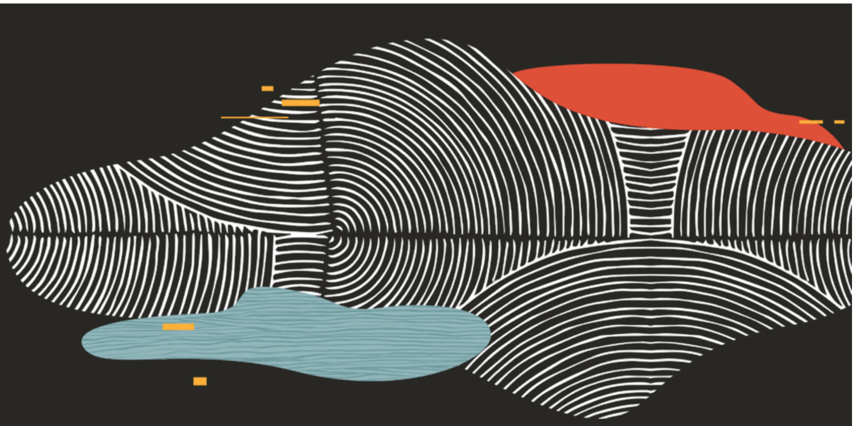
1 Types of Workloads

2 Challenges of Mixed Workloads

3 Ways to Manage Mixed Workloads

4 Demo

5 Lessons Learned



Agenda

1 Types of Workloads

2 Challenges of Mixed Workloads

3 Ways to Manage Mixed Workloads

4 Demo

5 Lessons Learned

Workload Characteristics

OLTP

- Many concurrent users
- SQL statements process a few rows at a time

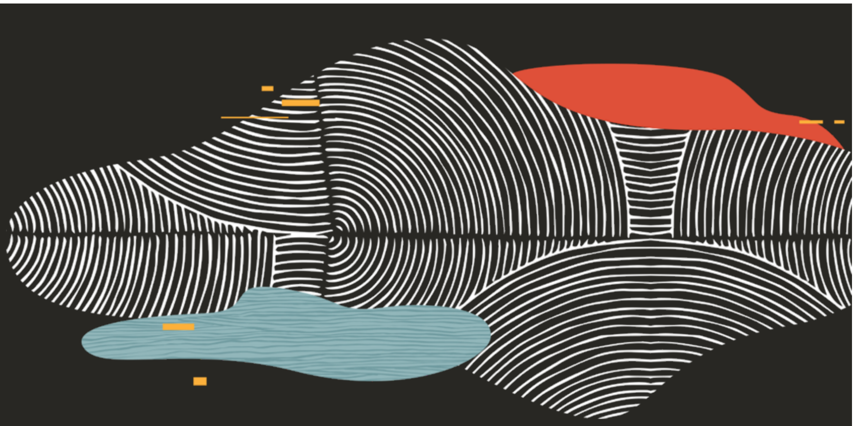
DW Queries

- Fewer concurrent users
- Data-intensive queries processing many rows

Data Loading and Processing

- Fewer concurrent processes
- DML processing many rows

Agenda



1 Types of Workloads

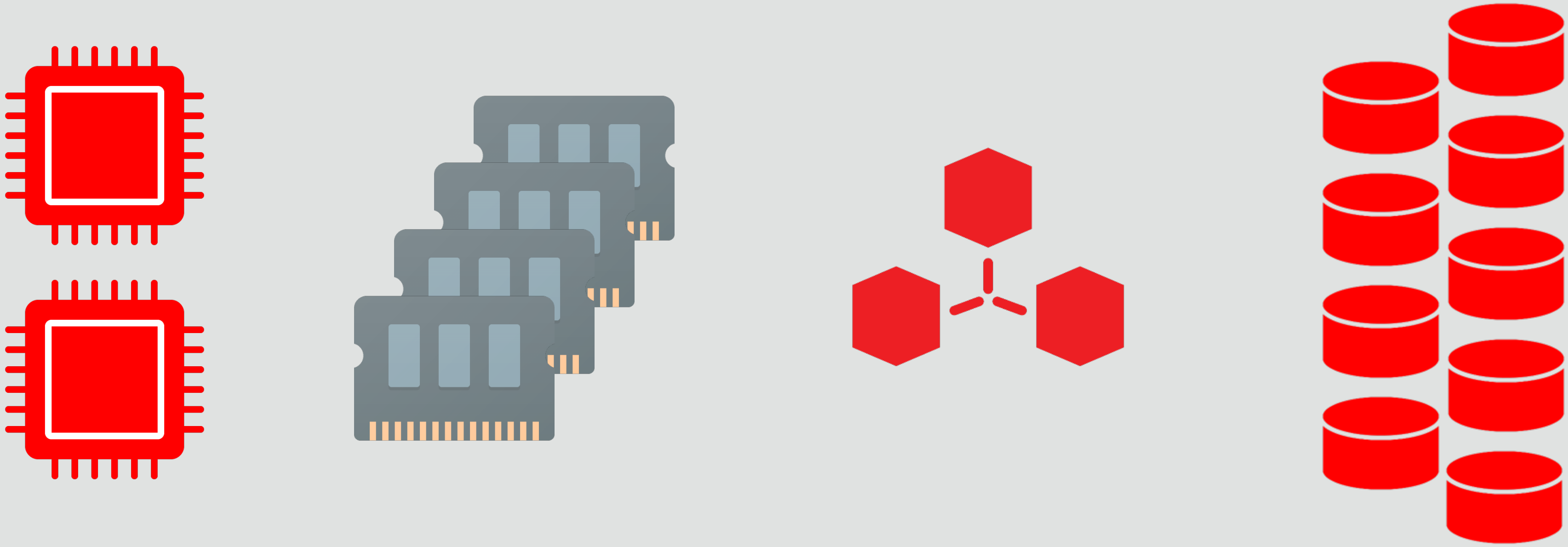
2 Challenges of Mixed Workloads

3 Ways to Manage Mixed Workloads

4 Demo

5 Lessons Learned

Competition for Resources



Memory

- Avoid competing for memory
- Competing for memory ends badly

ORACLE
REAL-WORLD PERFORMANCE

Network and Storage

- Oracle uses CPU when performing network and storage I/O
- Limiting CPU naturally limits network and storage I/O
- Focus on CPU

ORACLE
REAL-WORLD PERFORMANCE

CPU

- The Oracle database uses a process based architecture: when you connect, a dedicated foreground process is started to serve your calls
- To perform efficiently, a process:

1. needs to get on CPU as quickly as possible
2. needs to stay on CPU as long as possible:
3. should minimize voluntary sleeps
4. should experience as few involuntary sleeps as possible

CPU Resources and OLTP Workloads

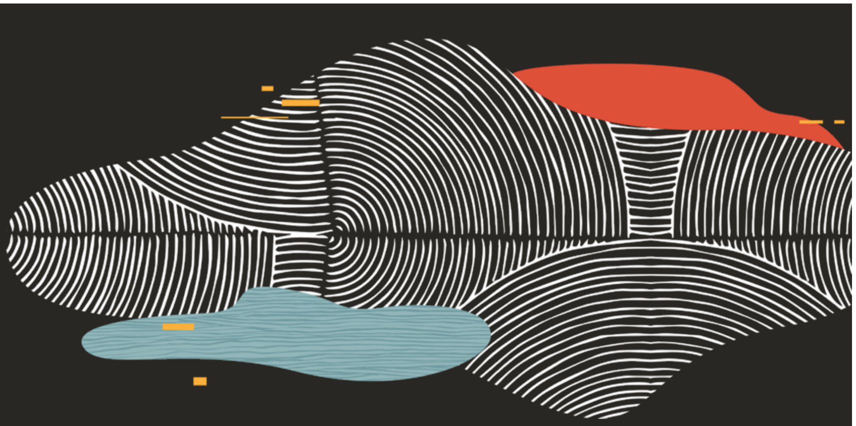
- As CPU Utilization increases, the chance of a process getting scheduled on CPU decreases
- This has a noticeable impact on OLTP performance at 60-70% CPU utilization

CPU Utilization	Chance of getting scheduled
50%	1 in 2
66%	1 in 3
80%	1 in 5
90%	1 in 10

The Mixed Workload Dilemma—Opposing resource management goals

Workload	Goal	CPU Strategy
OLTP	Fast Response Time	Minimize
Analytical Queries	Throughput and Response Time	Maximize
Data Processing	Throughput	Maximize

Agenda



1 Types of Workloads

2 Challenges of Mixed Workloads

3 Ways to Manage Mixed Workloads

4 Demo

5 Lessons Learned

Ways to Manage Mixed Workloads

Multiple Databases

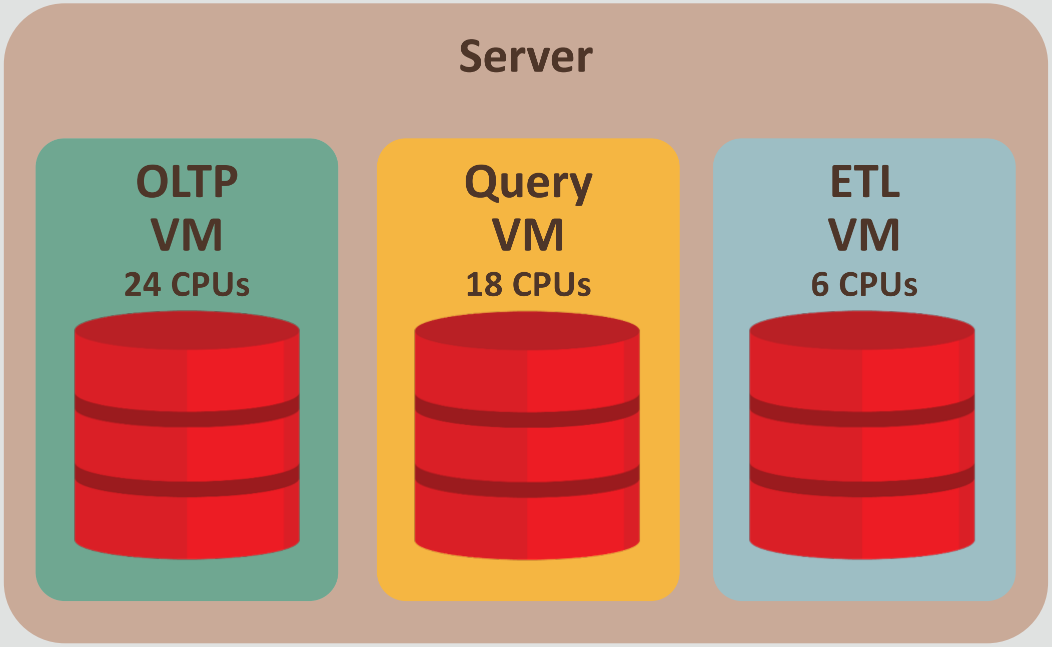
- Virtual Machines
- Instance Caging
- Multitenant

Single Database

- RAC Services
- Database Resource Manager

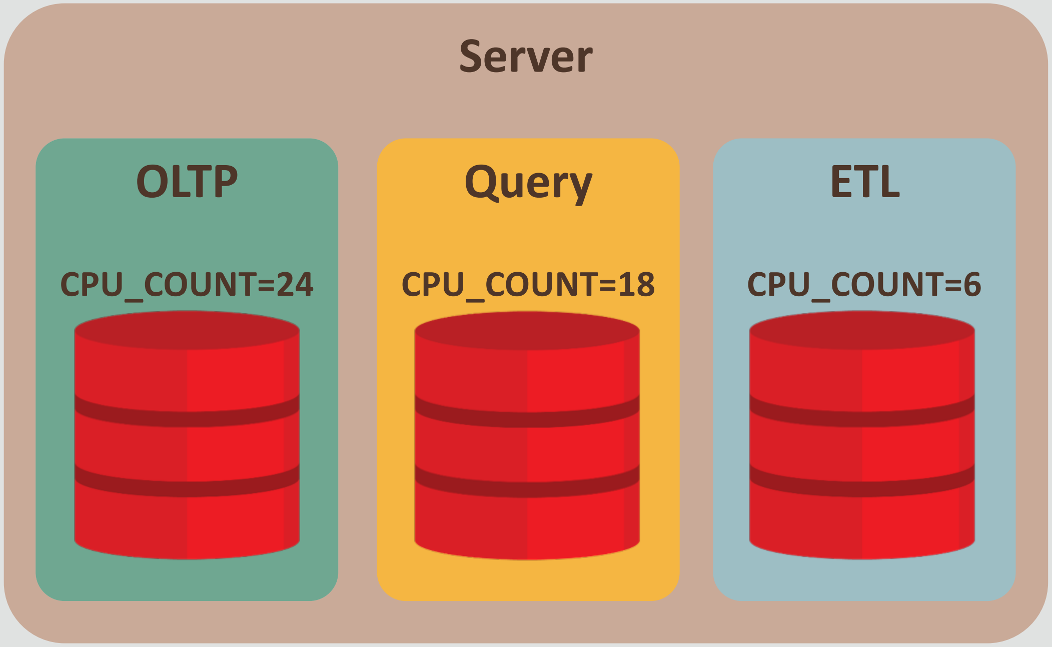
Multiple Databases: Virtual Machines

- Allocate virtual machines for each workload
- Workloads cannot use more than allocated CPUs



Multiple Databases: Instance Caging

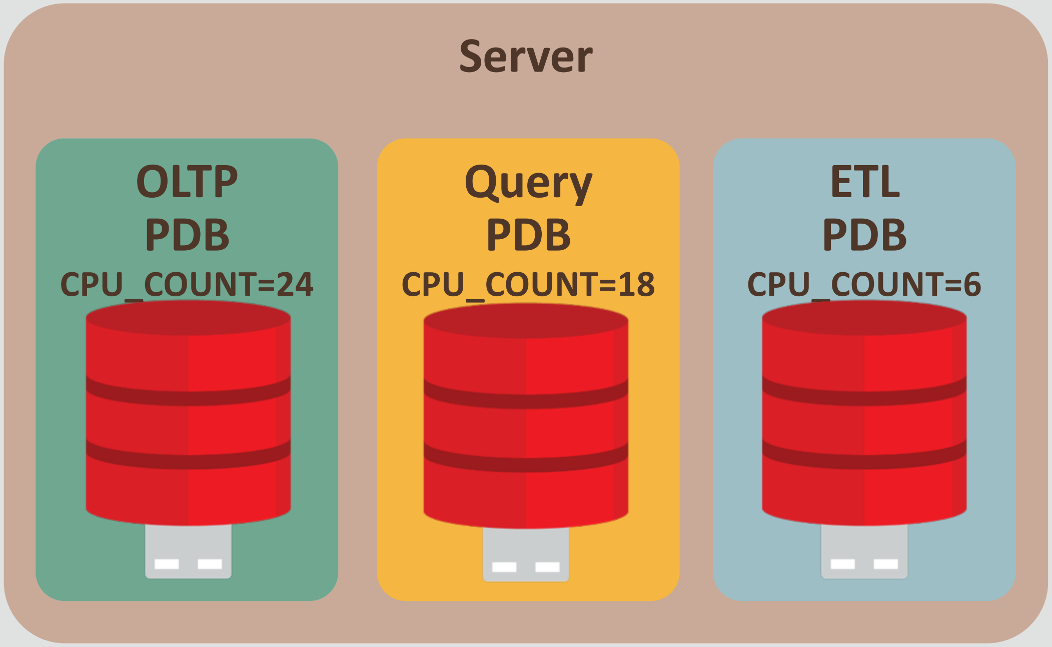
- Use instance caging to control the number of processes on CPU
- Use the CPU_COUNT parameter to control
- Enable a Database Resource Manager plan



ORACLE REAL-WORLD PERFORMANCE

Multiple Databases: Multitenant

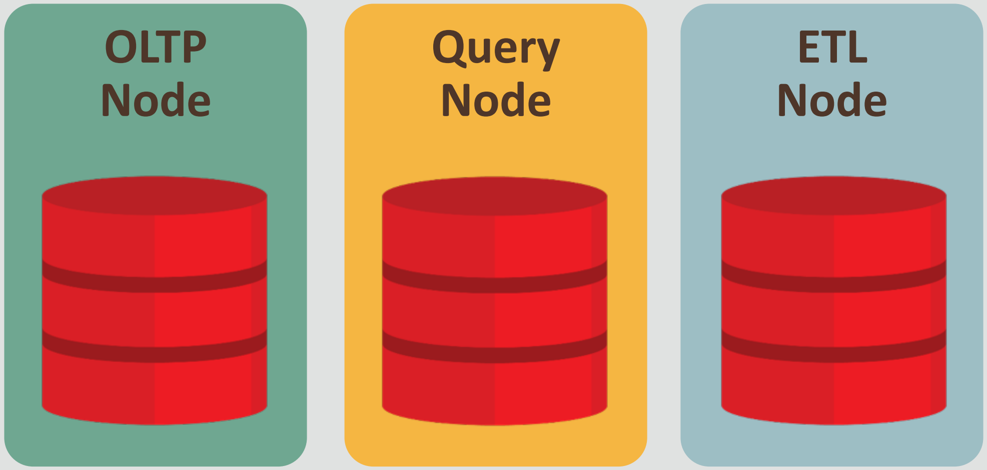
- Enable a CDB resource plan
- Use instance caging to limit CPU of individual PDBs
- Or use Shares or Limits



ORACLE
REAL-WORLD PERFORMANCE

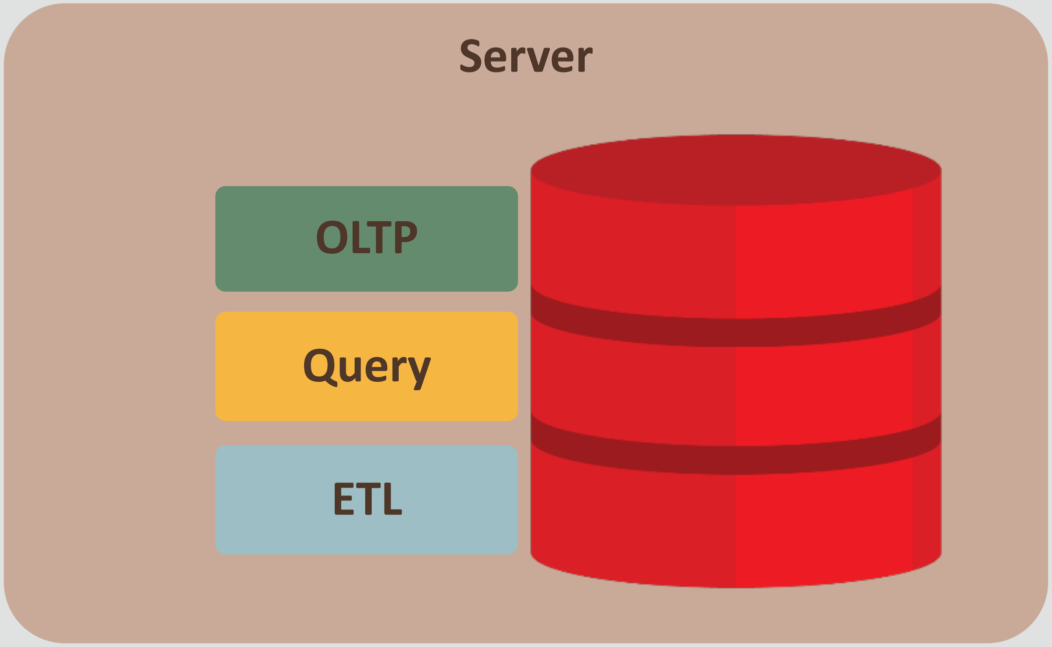
Single Database: RAC Services

- Use services for different workloads
- Map services for different workloads to different nodes in a cluster



Single Database: DBRM Consumer Groups

- Create a DBRM plan
- Map workloads to different consumer groups
- Use Shares or Limits



Database Resource Manager

Shares

Divide resources between workloads using ratios

Limits

Set hard limits on CPU utilization for each workload

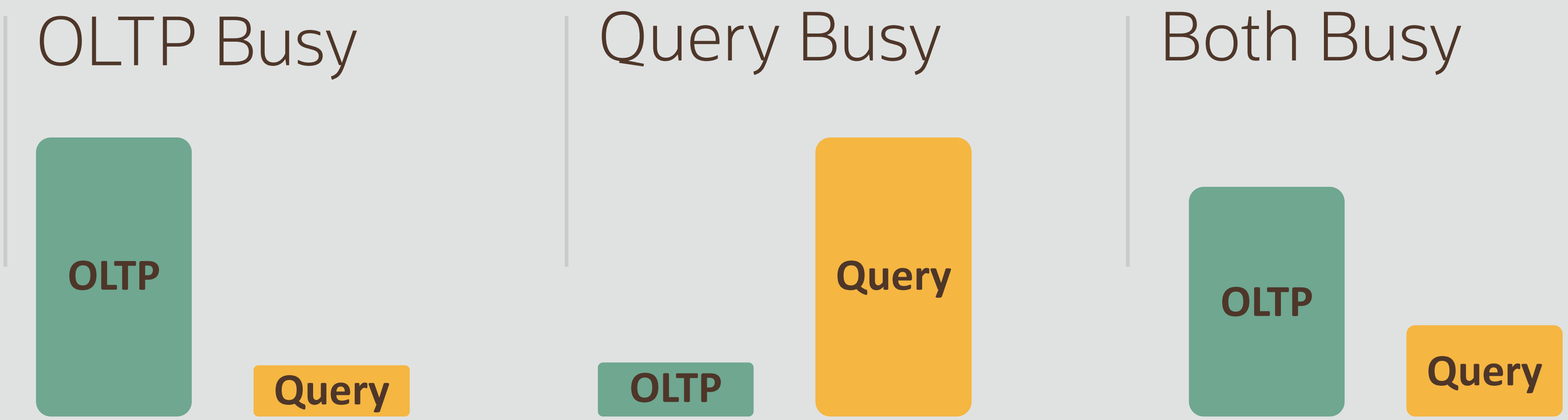
Parallel Queuing

Control the number of PX processes used by each workload

These can be combined to develop a more sophisticated CPU utilization strategy

Shares

Workload	Shares
OLTP	7
Queries	3
Total	10



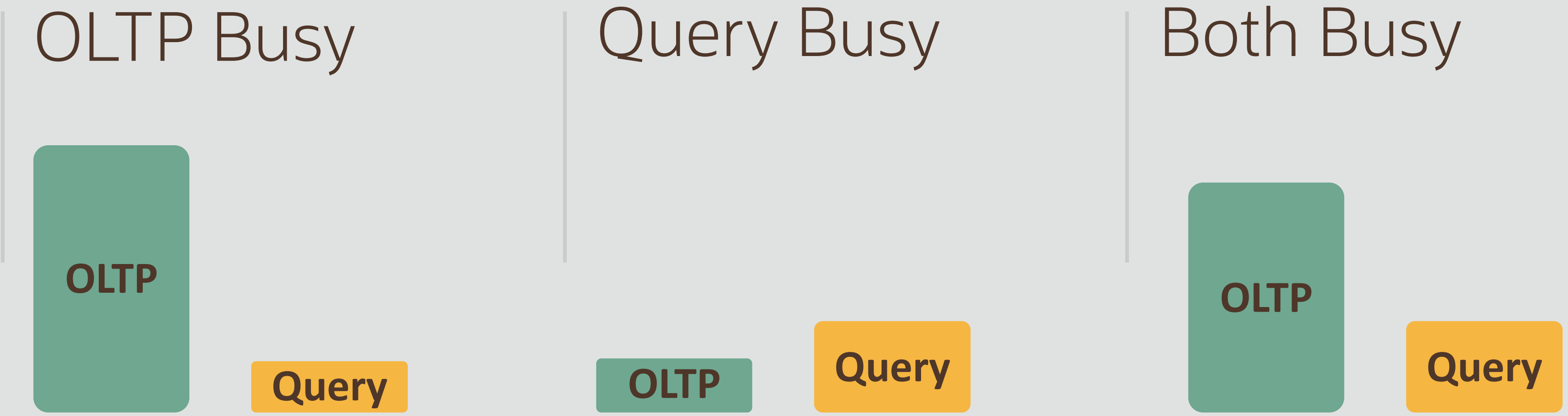
OLTP Busy

Query Busy

Both Busy

Limits

Workload	Limits
OLTP	
Queries	30%



OLTP Busy

OLTP

Query

Query Busy

OLTP

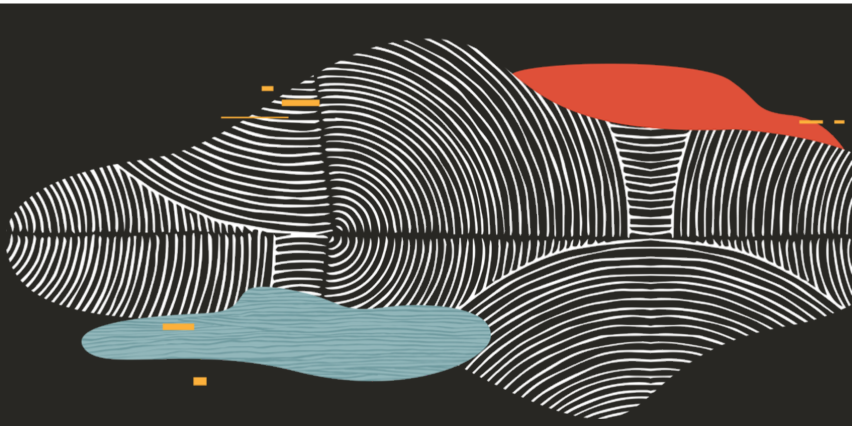
Query

Both Busy

OLTP

Query

Agenda



1 Types of Workloads

2 Challenges of Mixed Workloads

3 Ways to Manage Mixed Workloads

4 Demo

5 Lessons Learned

Demo Workloads

OLTP

- Lots of users playing online games
- Short transaction times
- Sensitive to high server CPU utilization
- This is the workload we want to protect

Queries

- 16 users running analytical queries which take a few seconds each
- Parallel execution with parallel degree 8
- Some queries perform table scans from disk and some scans are from memory

ETL

- Single user performing an ELT strategy
  1. Load Data
  2. Remove duplicates
  3. Transformations
  4. Aggregation
- Parallel execution with parallel degree 16

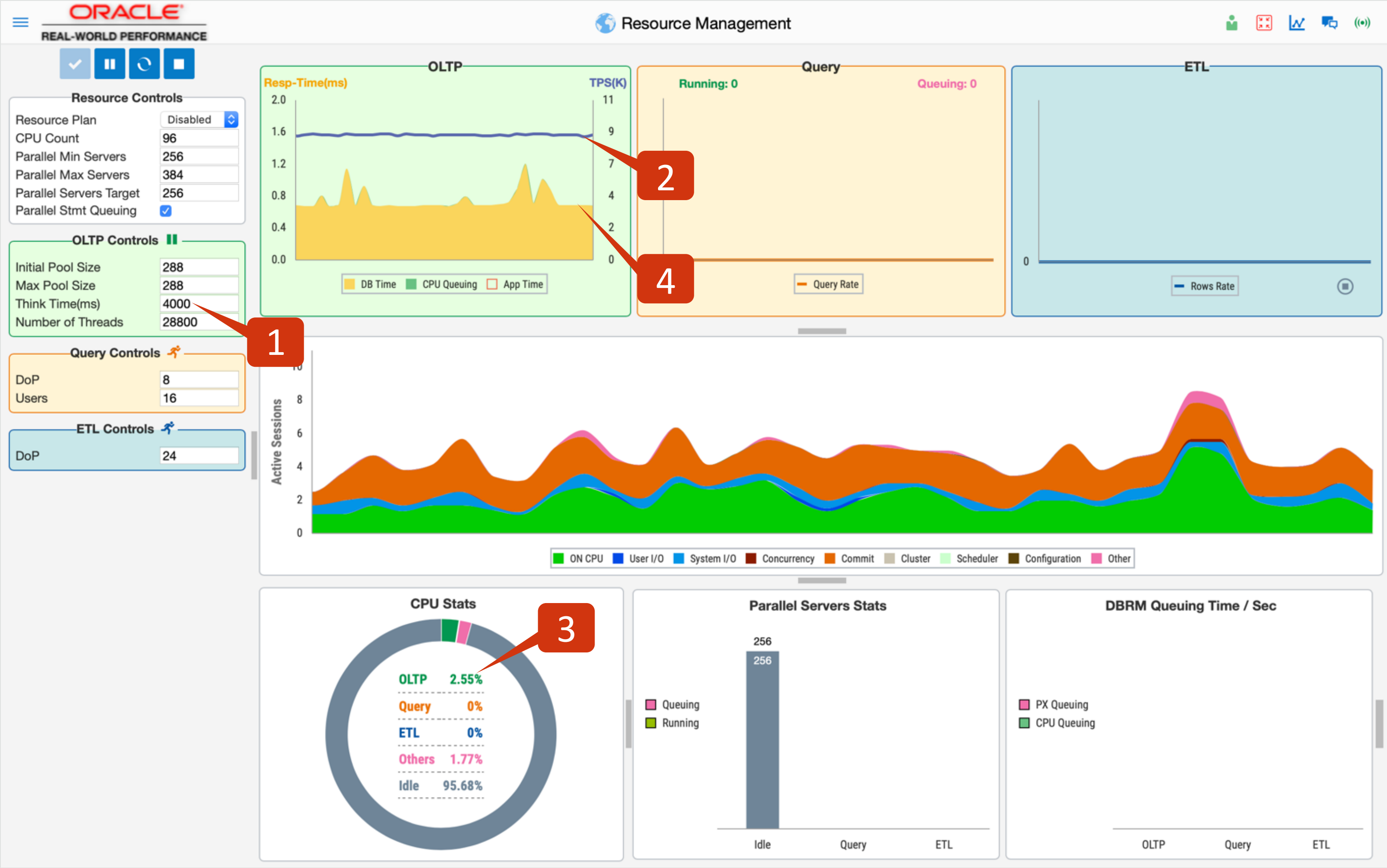
OLTP Workload

- We can control the workload by changing Think Time

  Decreasing Think Time increases demand
  Increasing Think Time decreases demand

- 4000ms represents low demand
- 500ms represents expected peak demand

1. Start the OLTP workload, setting Think Time to 4000 ms, representing low demand
2. About 9000 transactions per second with
3. OLTP CPU is about 3%
4. Sub-millisecond response time in the database



Resource Management

Resource Controls

Resource Plan	Disabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	4000
Number of Threads	28800

Query Controls

DoP	8
Users	16

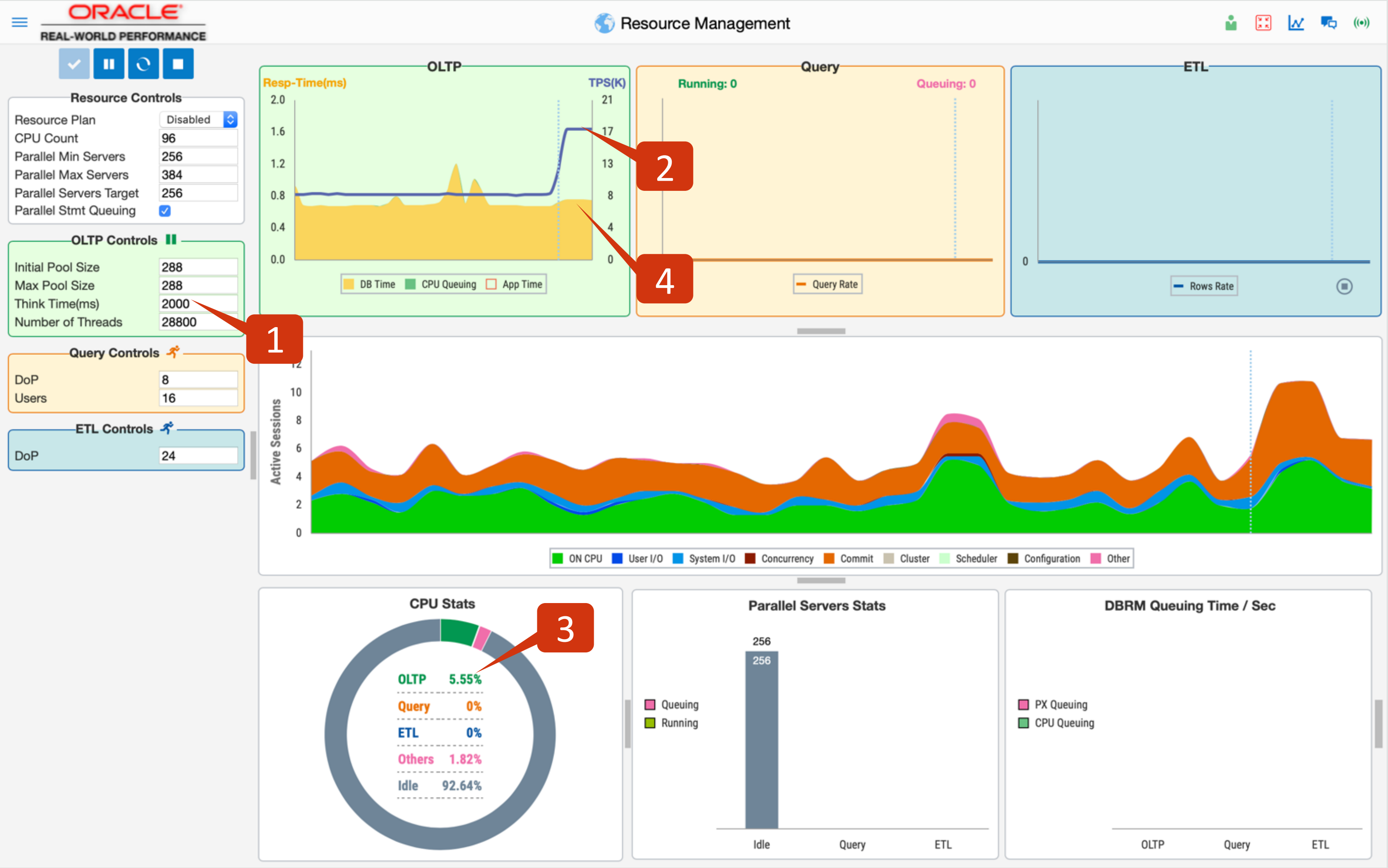
ETL Controls

DoP	24

CPU Stats

OLTP	2.55%
Query	0%
ETL	0%
Others	1.77%
Idle	95.68%

1. Adjust Think Time to 2000 ms, representing medium demand
2. TPS doubles
3. OLTP CPU doubles
4. Response time in the database is almost unchanged



Resource Controls

Resource Plan: Disabled
CPU Count: 96
Parallel Min Servers: 256
Parallel Max Servers: 384
Parallel Servers Target: 256
Parallel Stmt Queuing: ☑

OLTP Controls

Initial Pool Size: 288
Max Pool Size: 288
Think Time(ms): 2000
Number of Threads: 28800

Query Controls

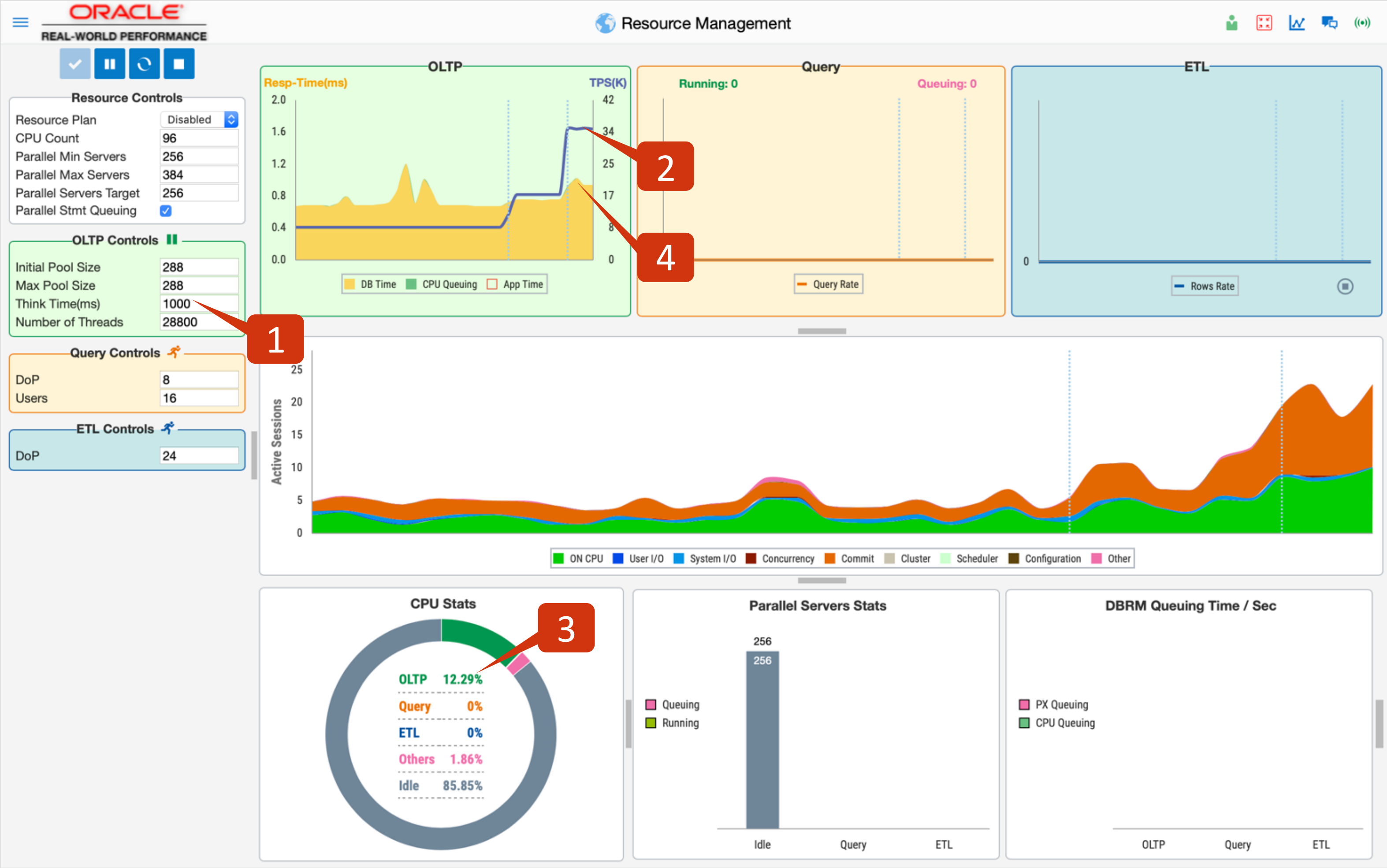
DoP: 8
Users: 16

ETL Controls

DoP: 24

CPU Stats	
OLTP	5.55%
Query	0%
ETL	0%
Others	1.82%
Idle	92.64%

1. Adjust Think Time to 1000 ms, representing high demand
2. TPS doubles
3. OLTP CPU doubles
4. Response time in the database remains sub-millisecond



ORACLE REAL-WORLD PERFORMANCE

Resource Management

Resource Controls

Resource Plan	Disabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	1000
Number of Threads	28800

Query Controls

DoP	8
Users	16

ETL Controls

DoP	24

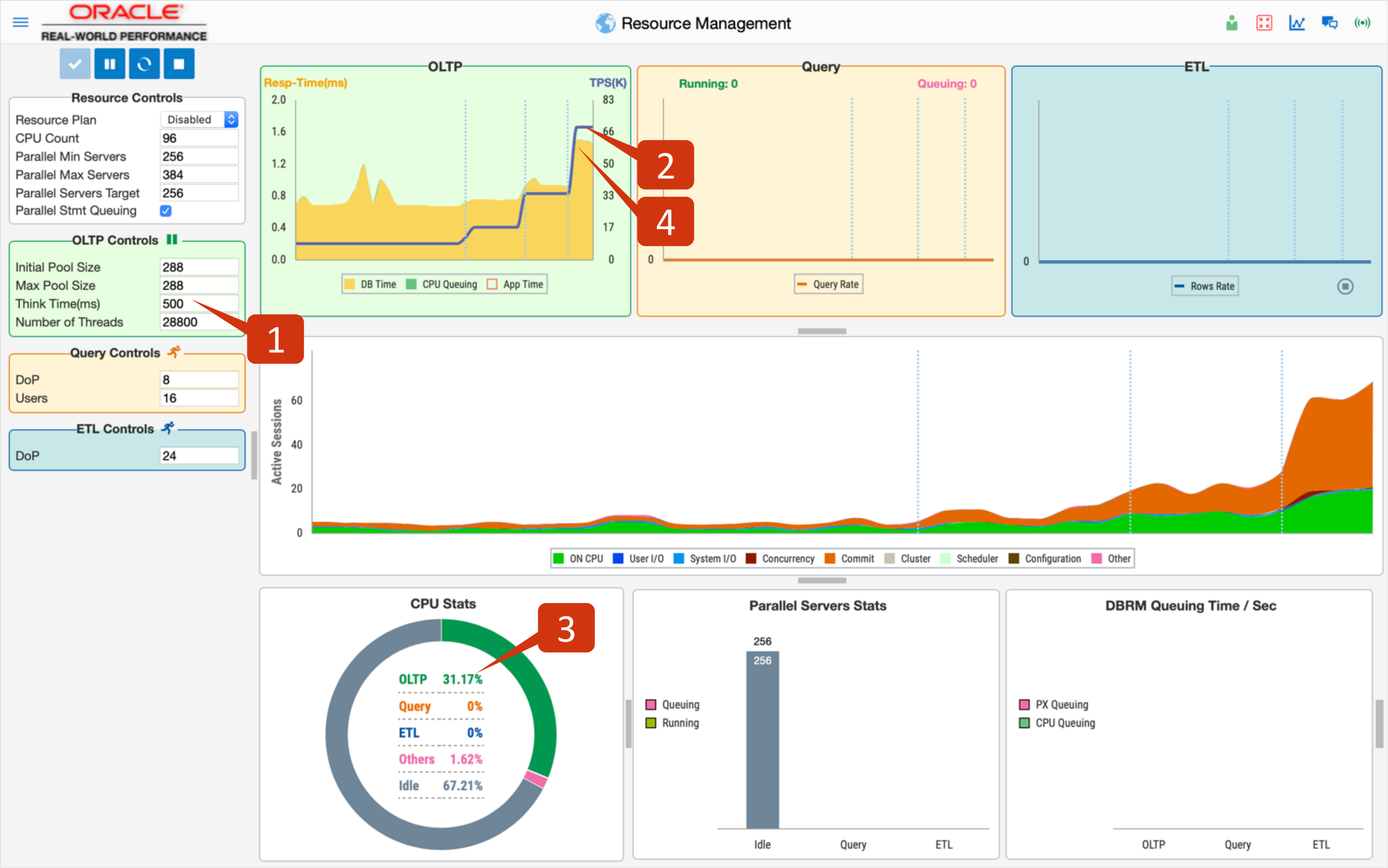
Query: Running: 0 Queuing: 0

CPU Stats

OLTP	12.29%
Query	0%
ETL	0%
Others	1.86%
Idle	85.85%

ORACLE REAL-WORLD PERFORMANCE

1. Adjust Think Time to 500 ms, representing peak demand
2. TPS doubles
3. OLTP CPU doubles
4. Response time in the database also increases



Resource Management

Resource Controls

Resource Plan: Disabled
CPU Count: 96
Parallel Min Servers: 256
Parallel Max Servers: 384
Parallel Servers Target: 256
Parallel Stmt Queuing: ☑

OLTP Controls

Initial Pool Size: 288
Max Pool Size: 288
Think Time(ms): 500
Number of Threads: 28800

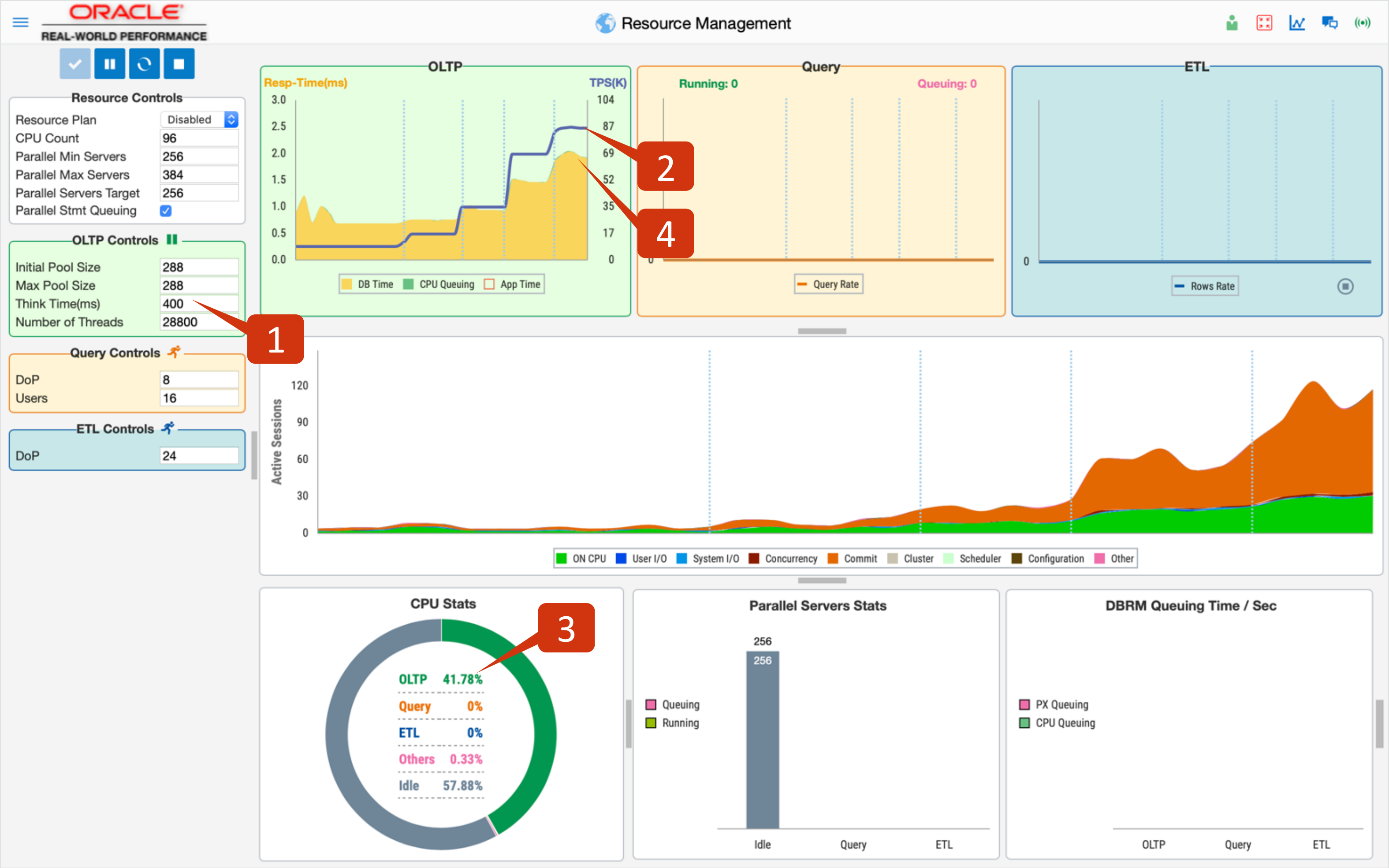
Query Controls

DoP: 8
Users: 16

ETL Controls

DoP: 24

1. Adjust Think Time to 400 ms to check for headroom
2. TPS increases
3. OLTP CPU increases
4. Response time in the database also increases



Resource Management

Resource Controls

Resource Plan	Disabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	400
Number of Threads	28800

Query Controls

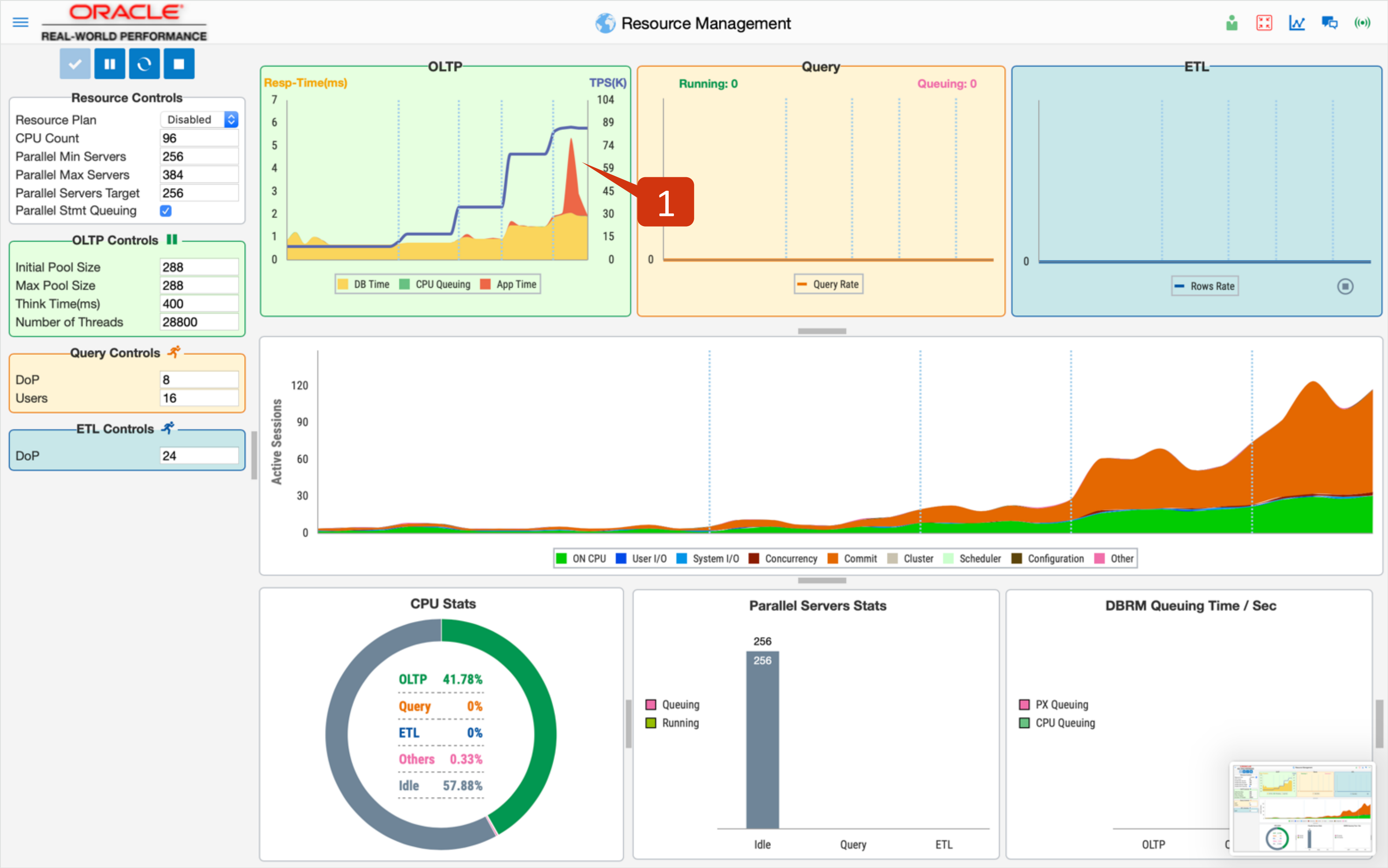
DoP	8
Users	16

ETL Controls

DoP	24

ORACLE
REAL-WORLD PERFORMANCE

1. Application Time indicates response time in the database plus time waiting for a connection. There is a short spike in waiting for a connection when demand increases.



Resource Management

Resource Controls

Resource Plan	Disabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	400
Number of Threads	28800

Query Controls

DoP	8
Users	16

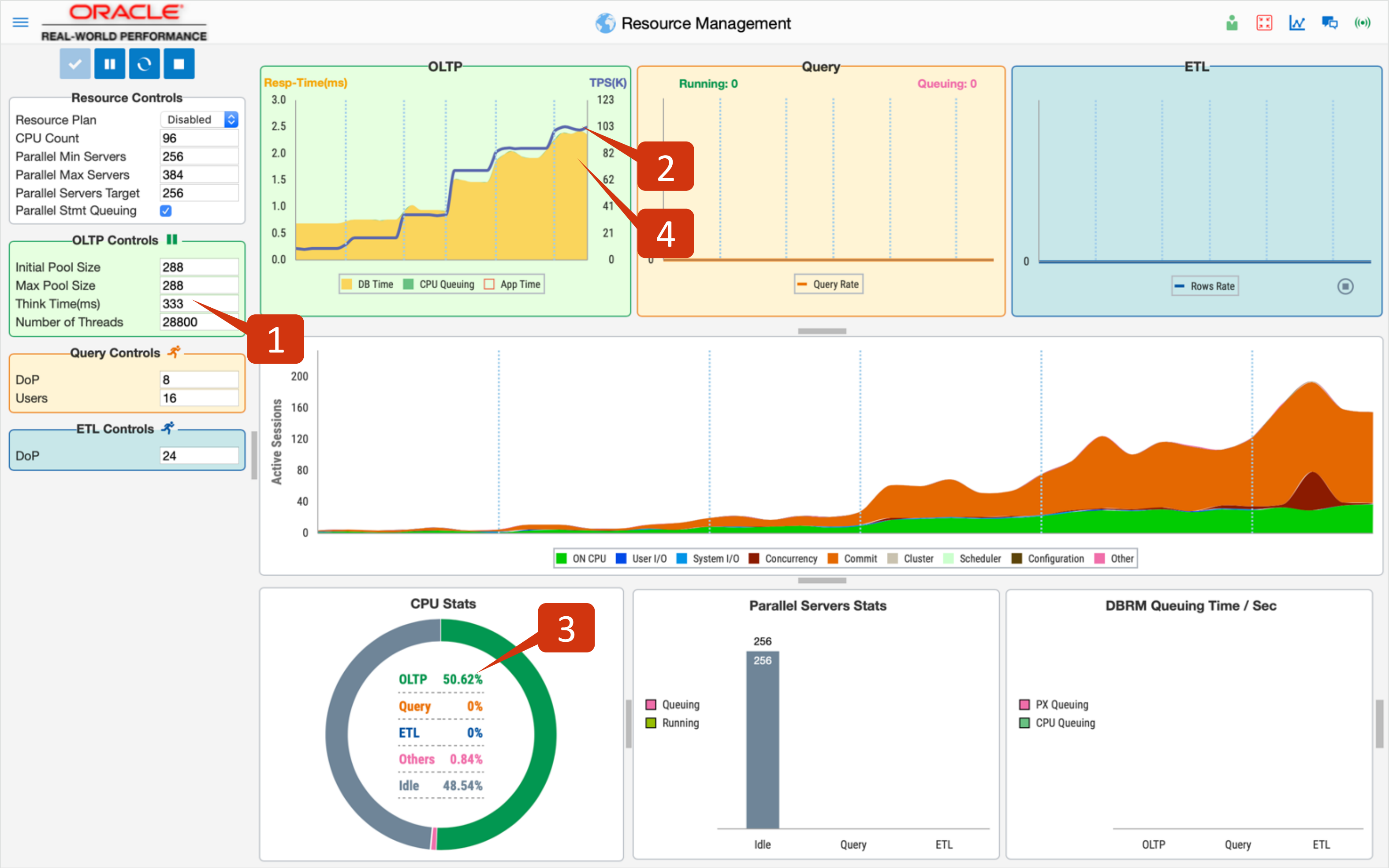
ETL Controls

DoP	24

Copyright © 2019 Oracle and/or its affiliates.

ORACLE
REAL-WORLD PERFORMANCE

1. Adjust Think Time to 333 ms to check for headroom
2. TPS increases
3. OLTP CPU increases
4. Response time in the database also increases



Resource Controls

Resource Plan	Disabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	333
Number of Threads	28800

Query Controls

DoP	8
Users	16

ETL Controls

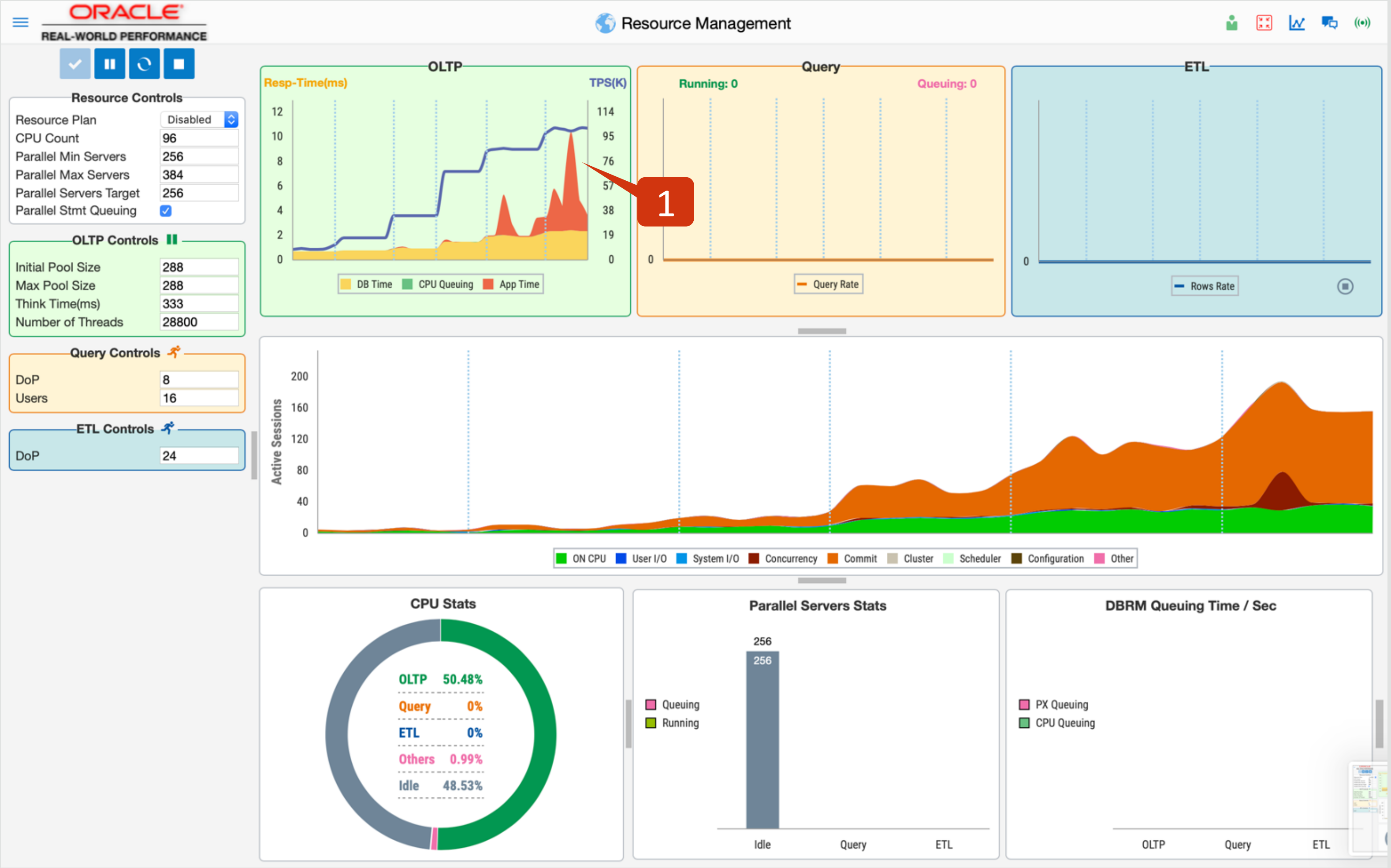
DoP	24

CPU Stats

OLTP	50.62%
Query	0%
ETL	0%
Others	0.84%
Idle	48.54%

ORACLE
REAL-WORLD PERFORMANCE

1. Some waiting for connections is now present at all times, degrading overall response times



Resource Controls

Resource Plan	Disabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	333
Number of Threads	28800

Query Controls

DoP	8
Users	16

ETL Controls

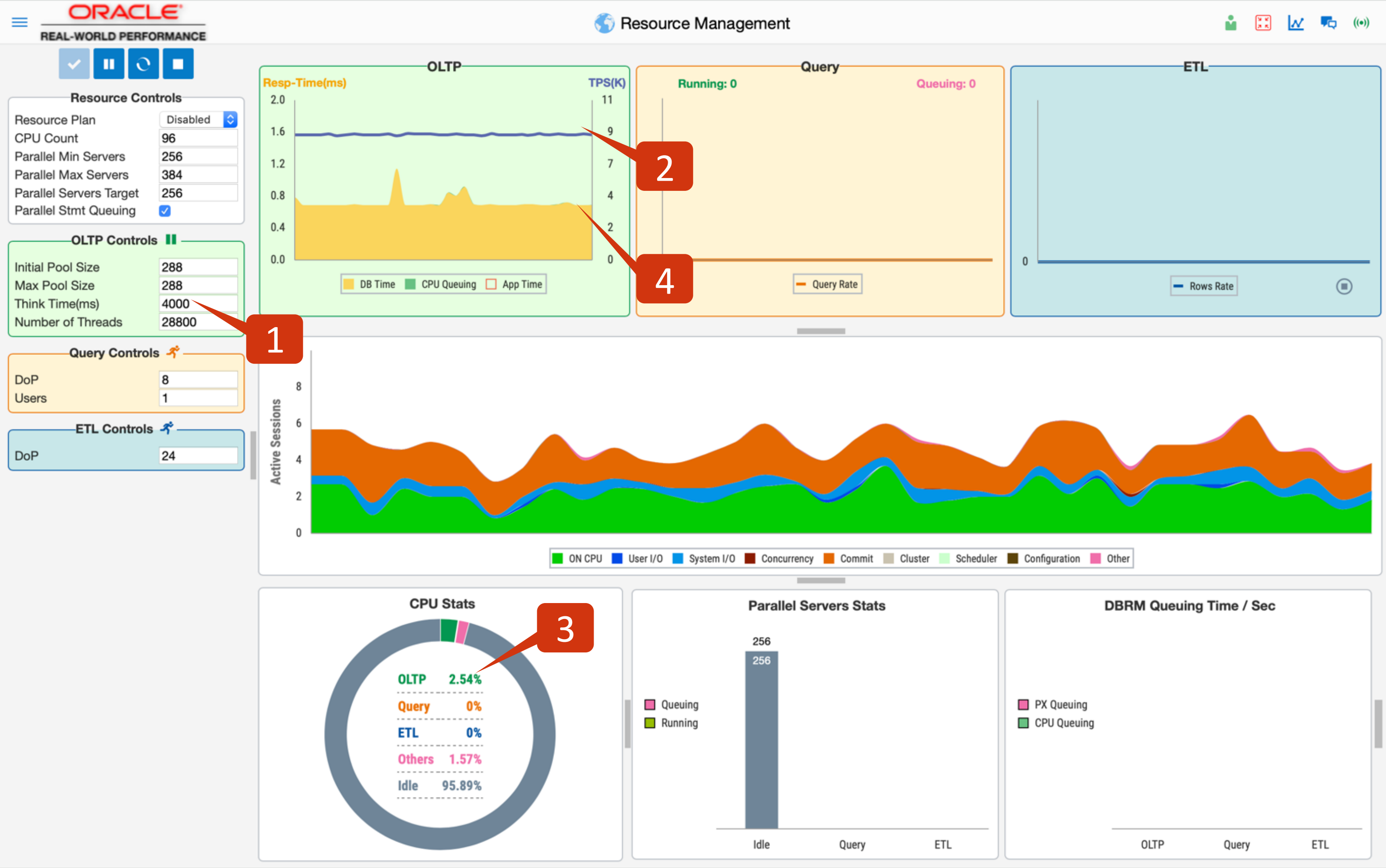
DoP	24

Query Workload

- We can control the workload by changing Users

  Increasing Users increases demand
  Decreasing Users decreases demand
- 1 User represents low demand
- 8 Users represents expected peak demand

1. Reset Think Time to 4000 ms to represent low demand
2. About 9000 transactions per second with
3. OLTP CPU is about 3%
4. Sub-millisecond response time in the database



Resource Management

Resource Controls

Resource Plan	Disabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	4000
Number of Threads	28800

Query Controls

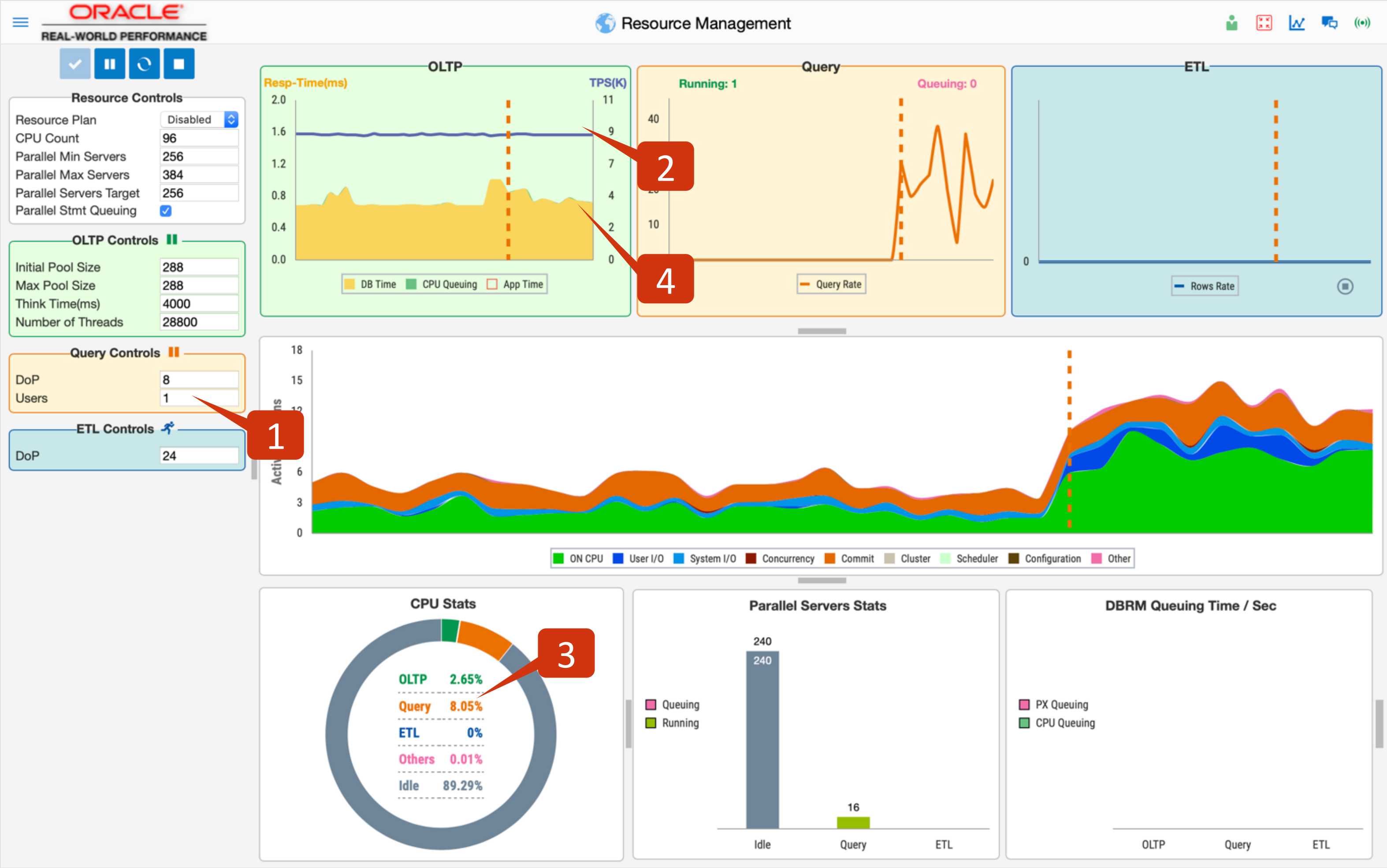
DoP	8
Users	1

ETL Controls

DoP	24

ORACLE
REAL-WORLD PERFORMANCE

1. Start the Query workload, setting Users to 1, representing low demand
2. OLTP throughput is unchanged
3. Query CPU is around 10%
4. Sub-millisecond response time in the database



ORACLE REAL-WORLD PERFORMANCE

Resource Management

Resource Controls

Resource Plan	Disabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	4000
Number of Threads	28800

Query Controls

DoP	8
Users	1

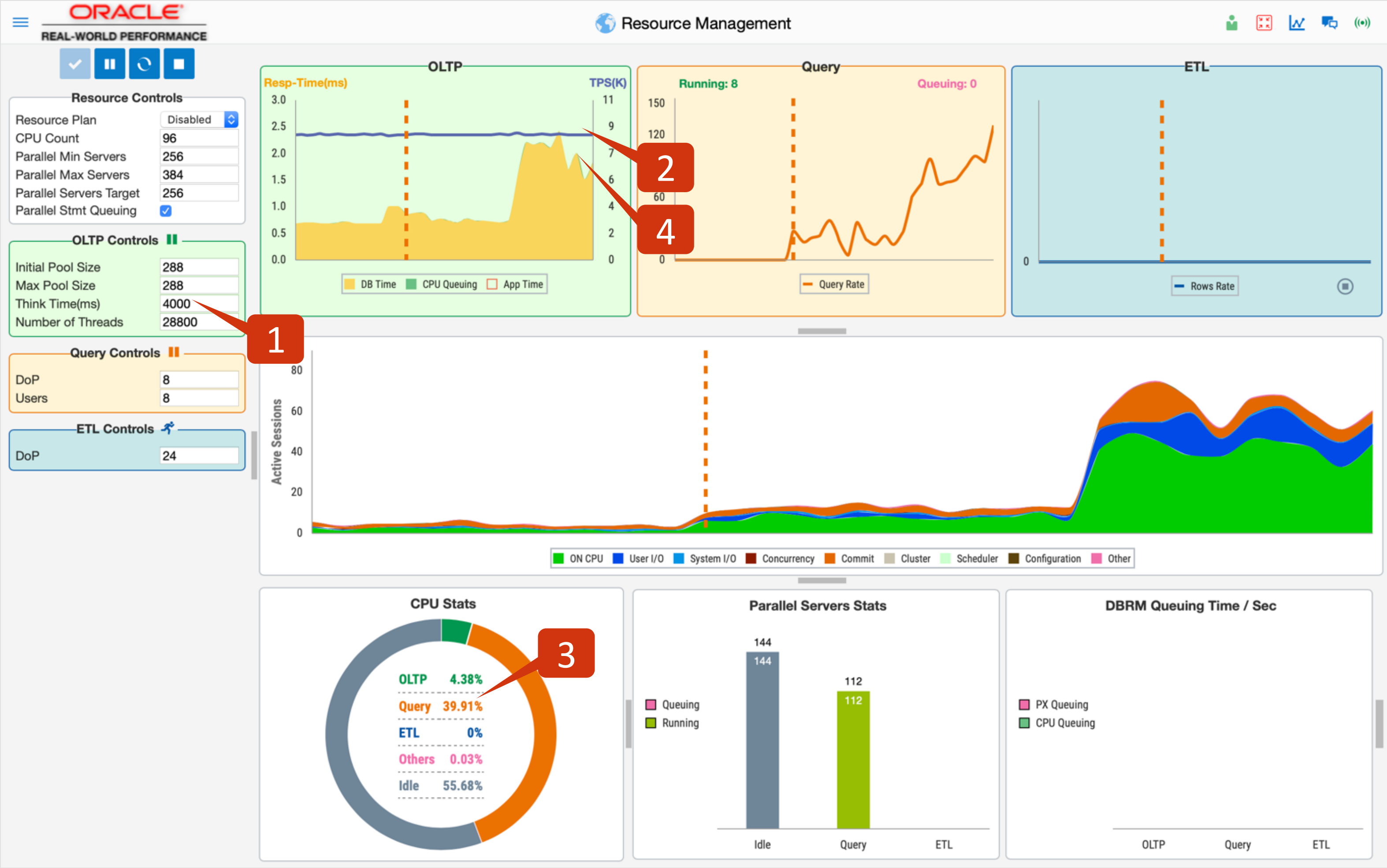
ETL Controls

DoP	24

CPU Stats

OLTP	2.65%
Query	8.05%
ETL	0%
Others	0.01%
Idle	89.29%

1. Adjust Users to 8, representing peak demand
2. OLTP throughput is unchanged
3. Query CPU is around 40%
4. Response time in the database increases



Resource Management

Resource Controls

Resource Plan	Disabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	4000
Number of Threads	28800

Query Controls

DoP	8
Users	8

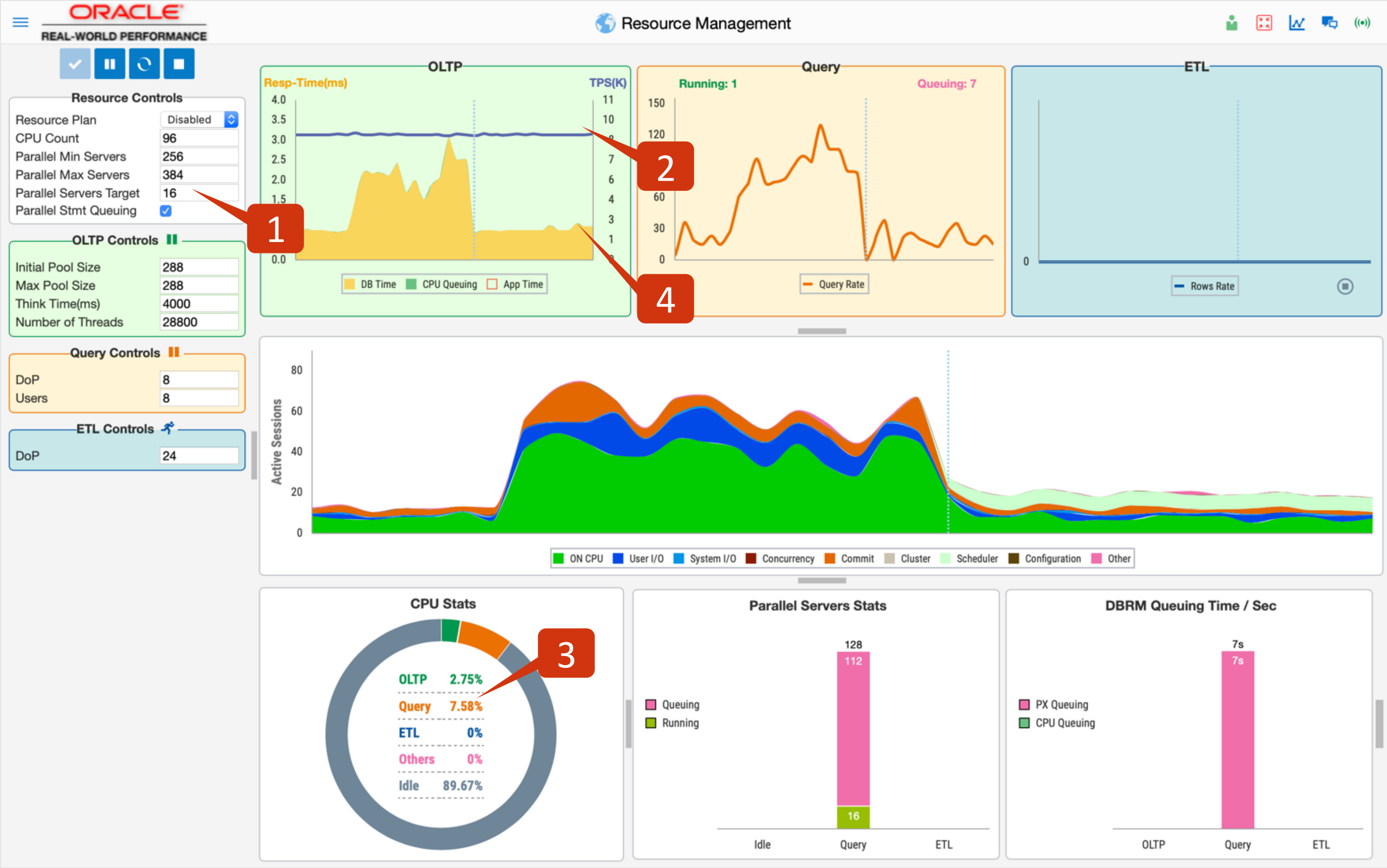
ETL Controls

DoP	24

CPU Stats

OLTP	4.38%
Query	39.91%
ETL	0%
Others	0.03%
Idle	55.68%

1. Adjust Parallel Servers Target to 16
2. OLTP throughput is unchanged
3. Query CPU is around 10%
4. Response time in the database improves



Resource Management

Resource Controls

Resource Plan	Disabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	16
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	4000
Number of Threads	28800

Query Controls

DoP	8
Users	8

ETL Controls

DoP	24

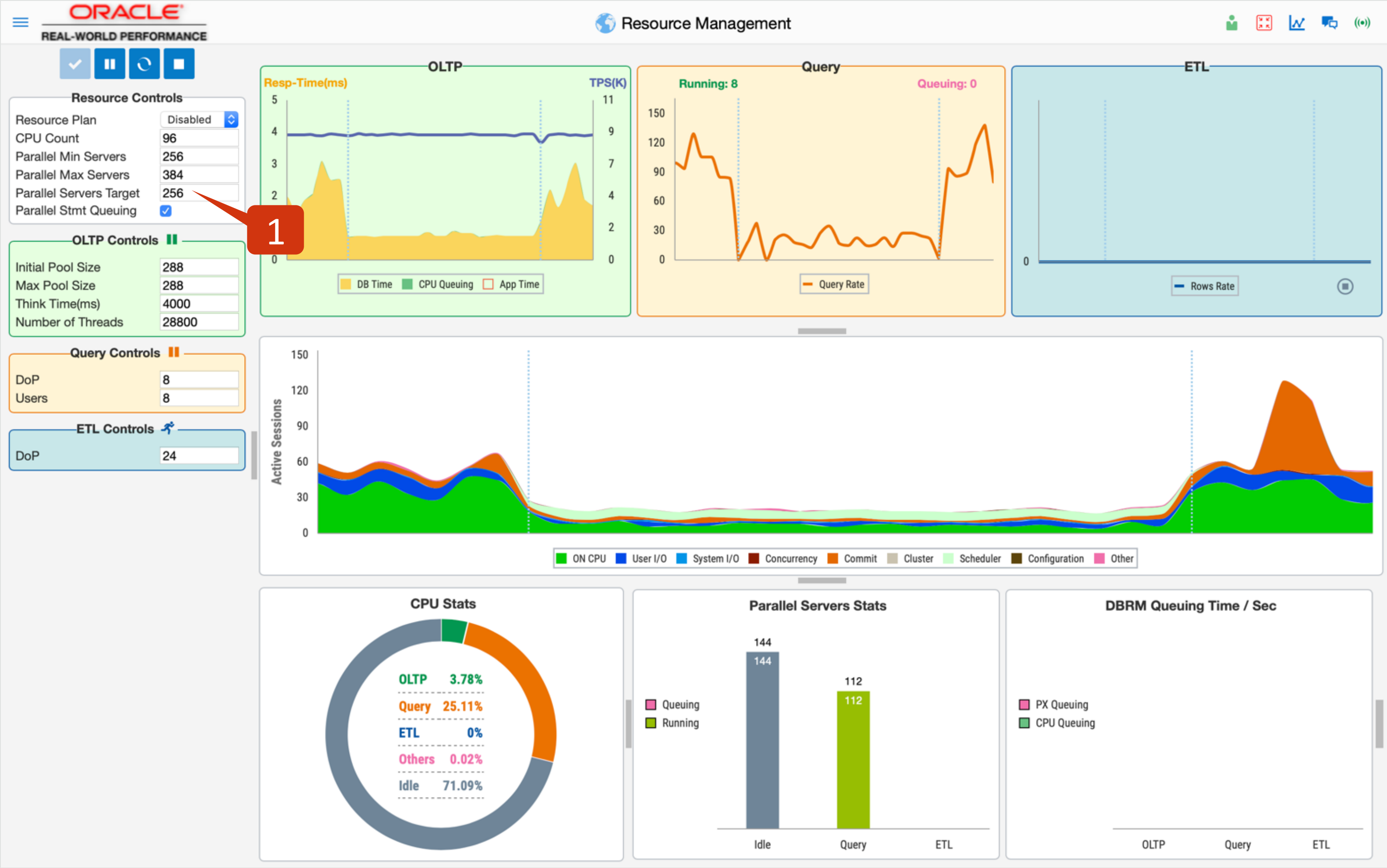
ETL Workload

- We can control the workload by changing Degree of Parallelism (DoP)

  Increasing DoP increases demand
  Decreasing DoP decreases demand

- Degree of Parallelism of 24 represents expected demand

1. Reset Parallel Servers Target to 256



Resource Controls

Resource Plan	Disabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	4000
Number of Threads	28800

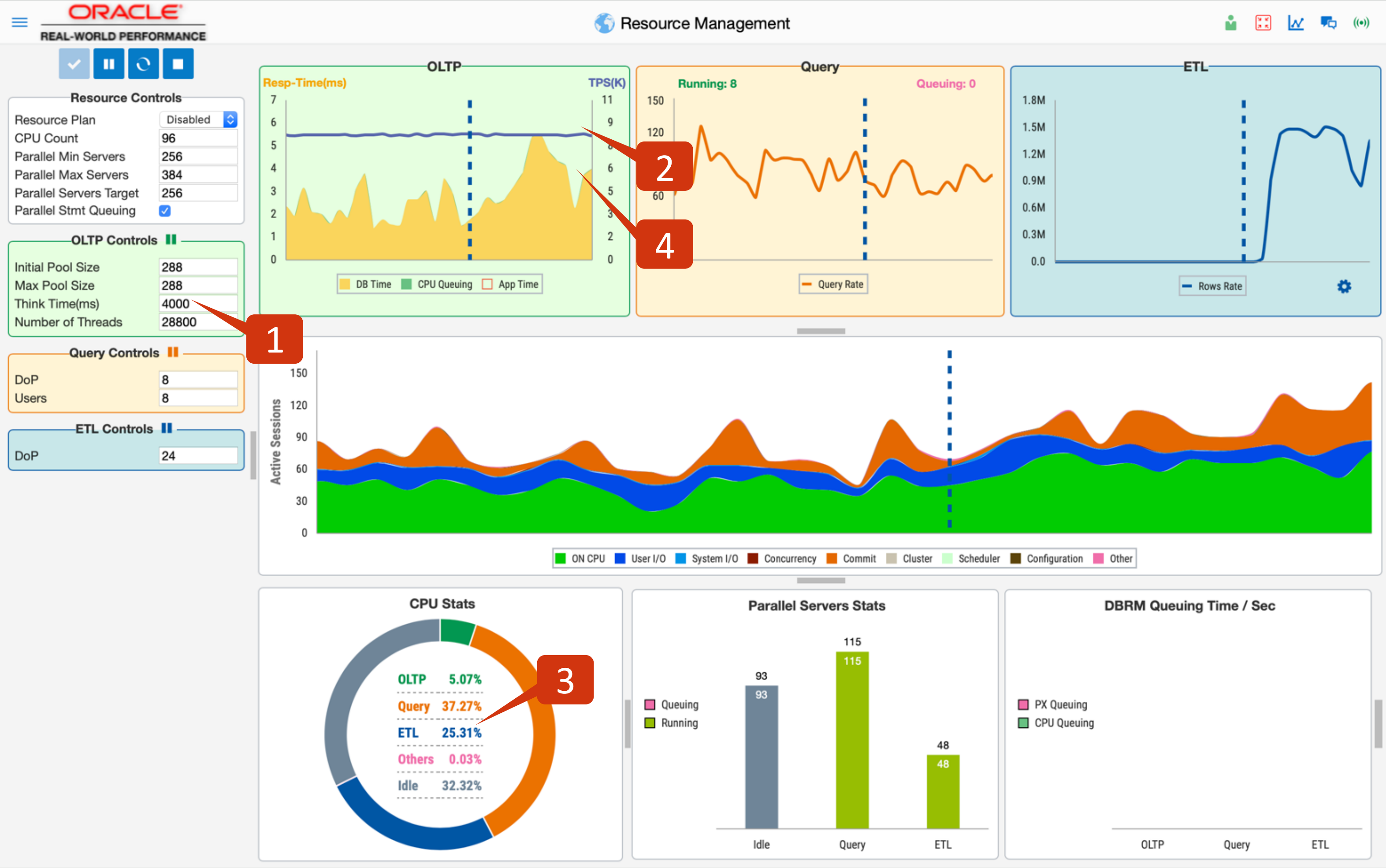
Query Controls

DoP	8
Users	8

ETL Controls

DoP	24

1. Start the ETL workload, setting DoP to 24, representing expected demand
2. OLTP throughput is unchanged
3. ETL CPU is around 25%
4. Response time in the database increases



Resource Management

Resource Controls

Resource Plan	Disabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	4000
Number of Threads	28800

Query Controls

DoP	8
Users	8

ETL Controls

DoP	24

CPU Stats

OLTP	5.07%
Query	37.27%
ETL	25.31%
Others	0.03%
Idle	32.32%

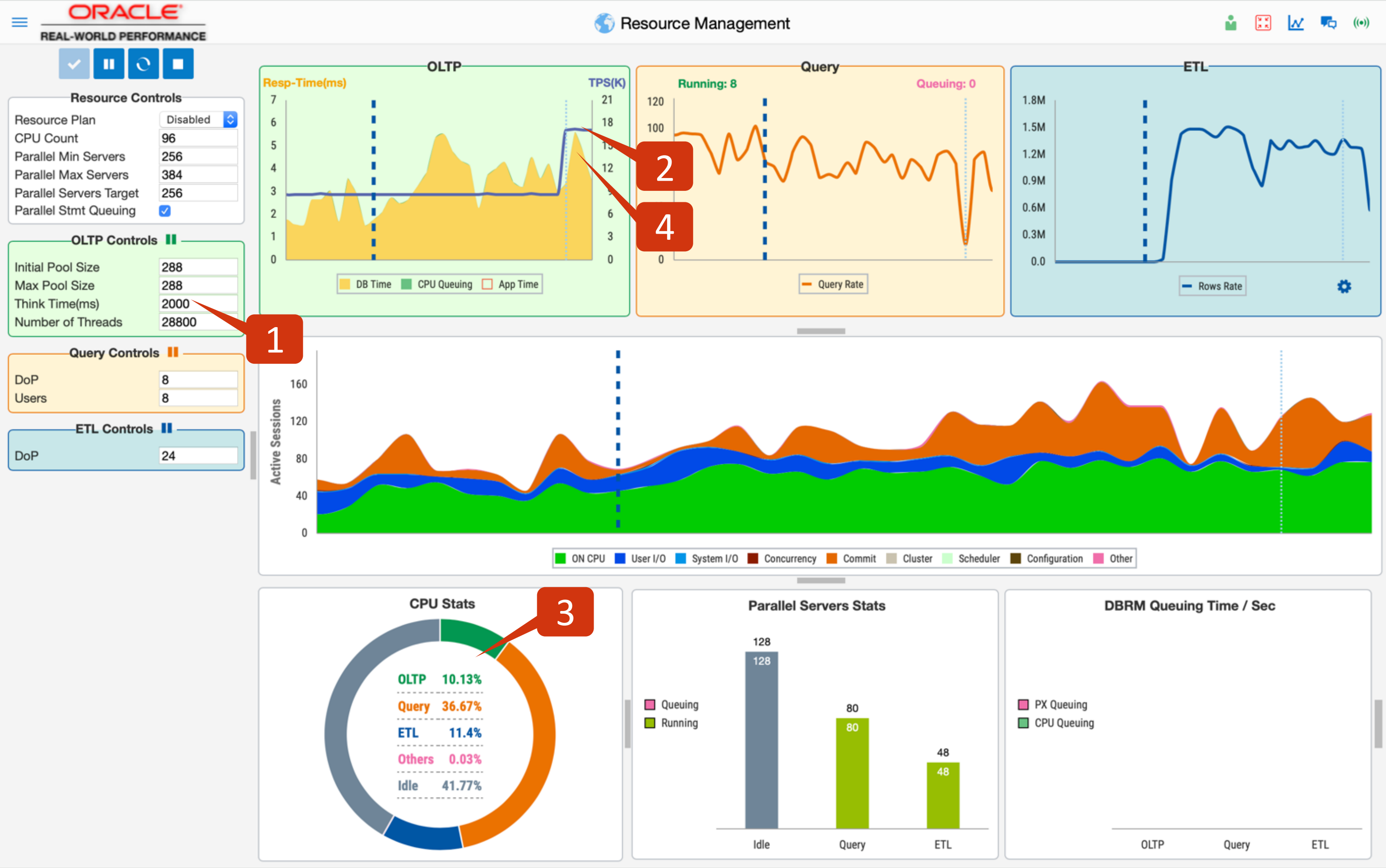
Increasing The Demand For OLTP

- We can control the OLTP workload by changing Think Time

  Decreasing Think Time increases demand
  Increasing Think Time decreases demand

- 4000ms represents low demand
- 500ms represents expected peak demand

1. Adjust Think Time to 2000 ms, representing medium demand
2. TPS doubles
3. OLTP CPU doubles
4. Response time in the database increases



ORACLE REAL-WORLD PERFORMANCE

Resource Management

Resource Controls

Resource Plan	Disabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	2000
Number of Threads	28800

Query Controls

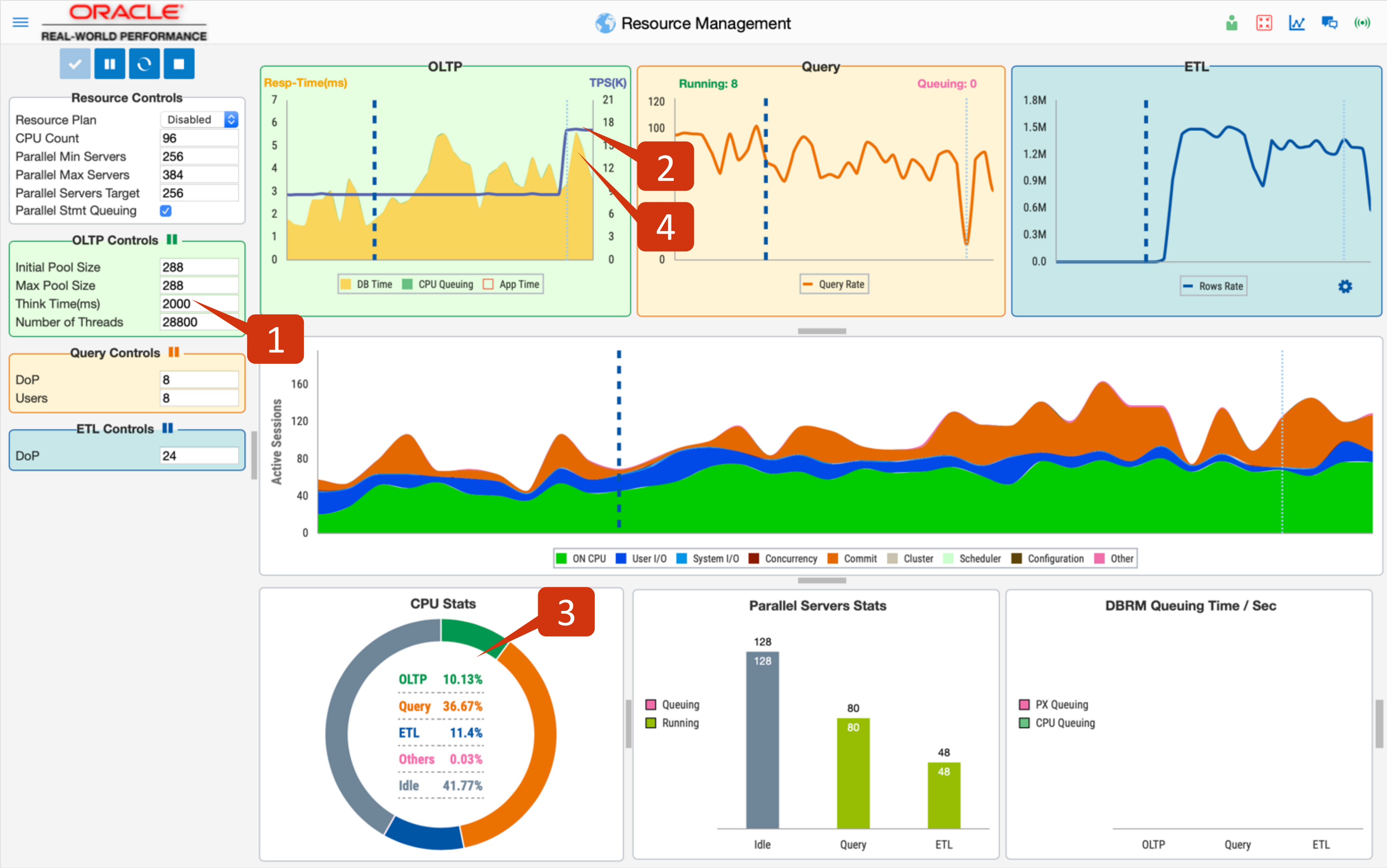
DoP	8
Users	8

ETL Controls

DoP	24

ORACLE REAL-WORLD PERFORMANCE

1. Adjust Think Time to 1000 ms, representing high demand
2. TPS doubles
3. OLTP CPU doubles
4. Response time in the database increases



Resource Management

Resource Controls

Resource Plan	Disabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	2000
Number of Threads	28800

Query Controls

DoP	8
Users	8

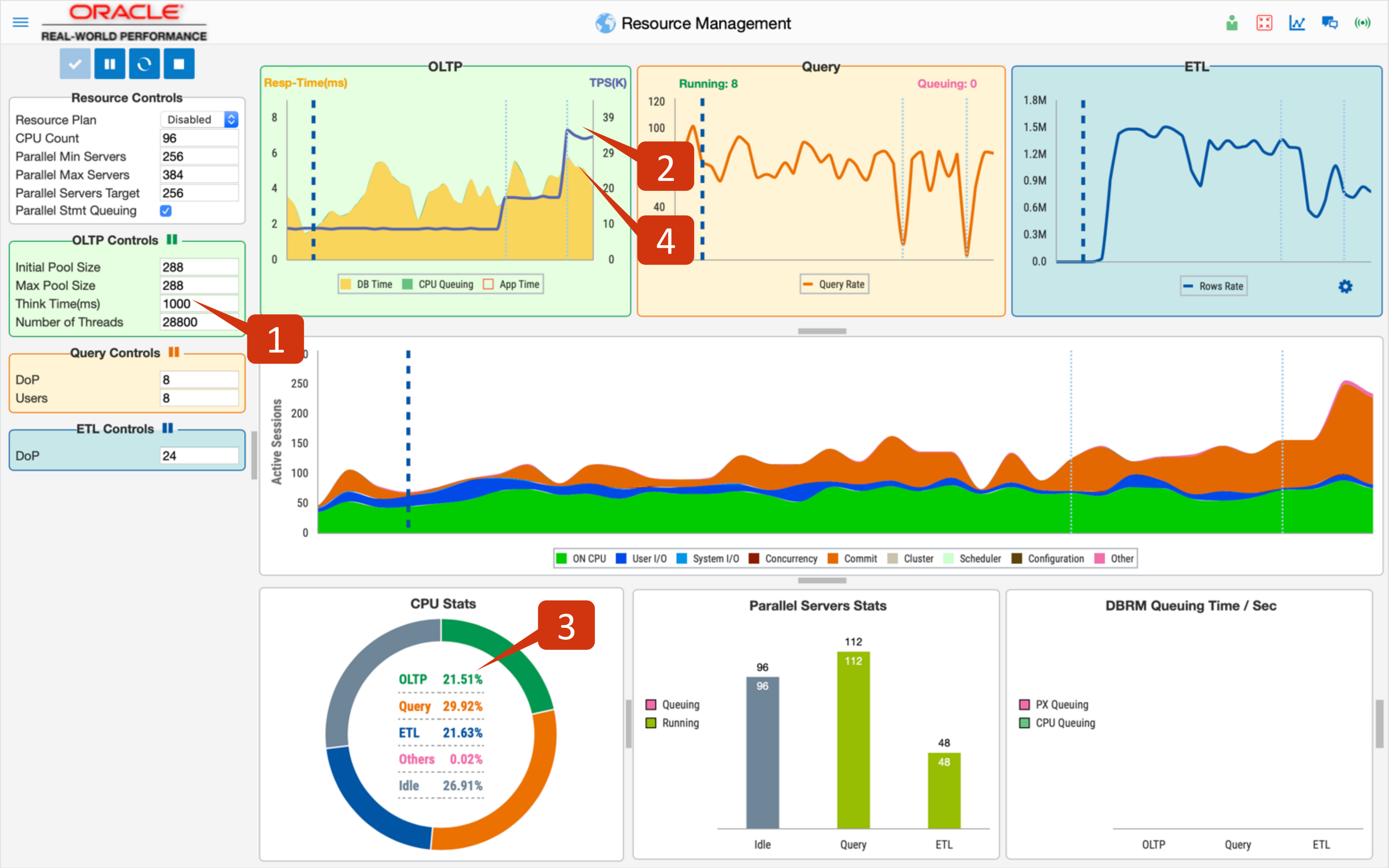
ETL Controls

DoP	24

CPU Stats

OLTP	10.13%
Query	36.67%
ETL	11.4%
Others	0.03%
Idle	41.77%

1. Adjust Think Time to 1000 ms, representing high demand
2. TPS doubles
3. OLTP CPU doubles
4. Response time in the database increases



Resource Management

Resource Controls

Resource Plan	Disabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	1000
Number of Threads	28800

Query Controls

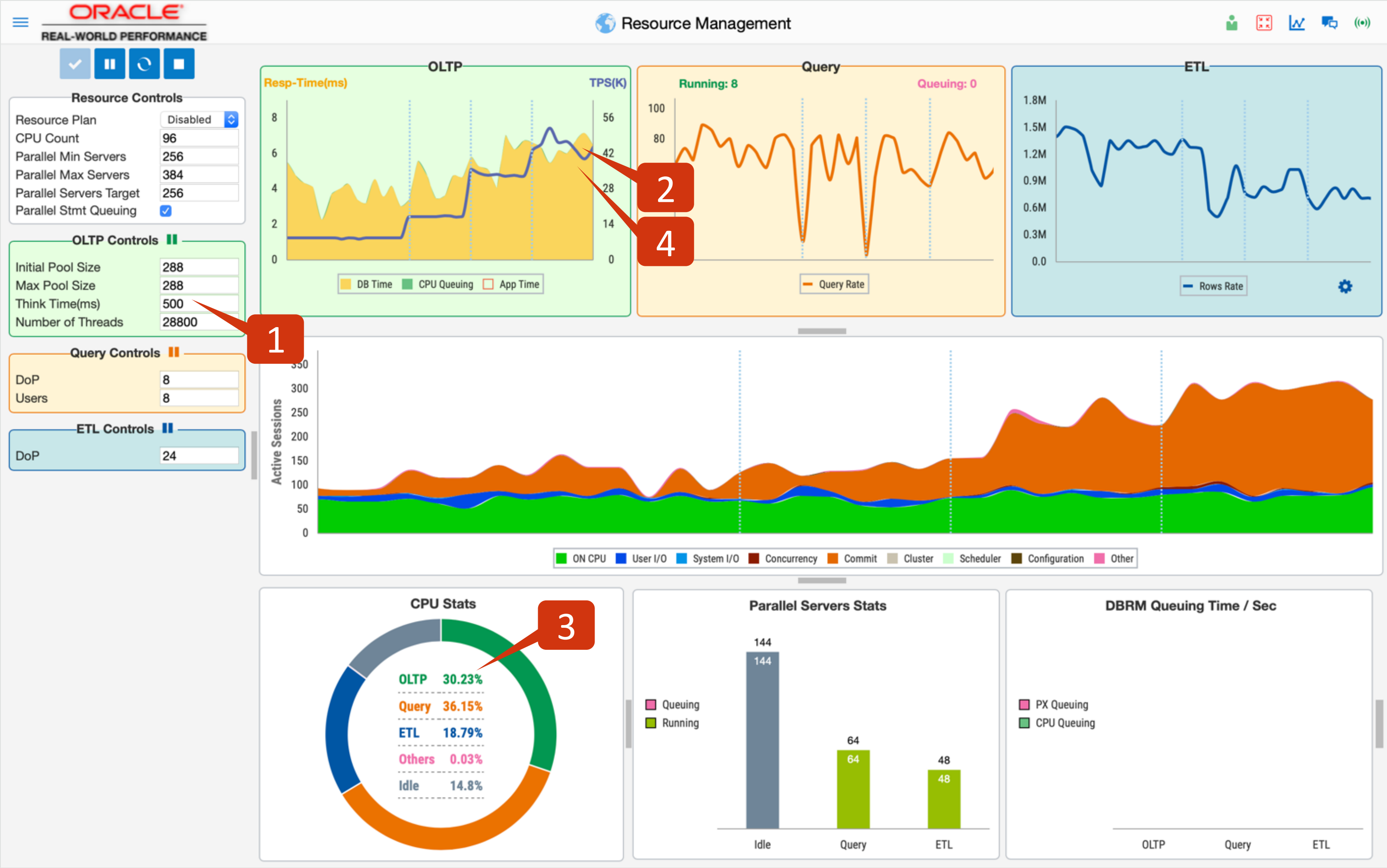
DoP	8
Users	8

ETL Controls

DoP	24

ORACLE
REAL-WORLD PERFORMANCE

1. Adjust Think Time to 500 ms, representing expected peak demand
2. TPS increases a little but becomes erratic
3. OLTP CPU increases
4. Response time in the database increases



Resource Management

Resource Controls

Resource Plan	Disabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	500
Number of Threads	28800

Query Controls

DoP	8
Users	8

ETL Controls

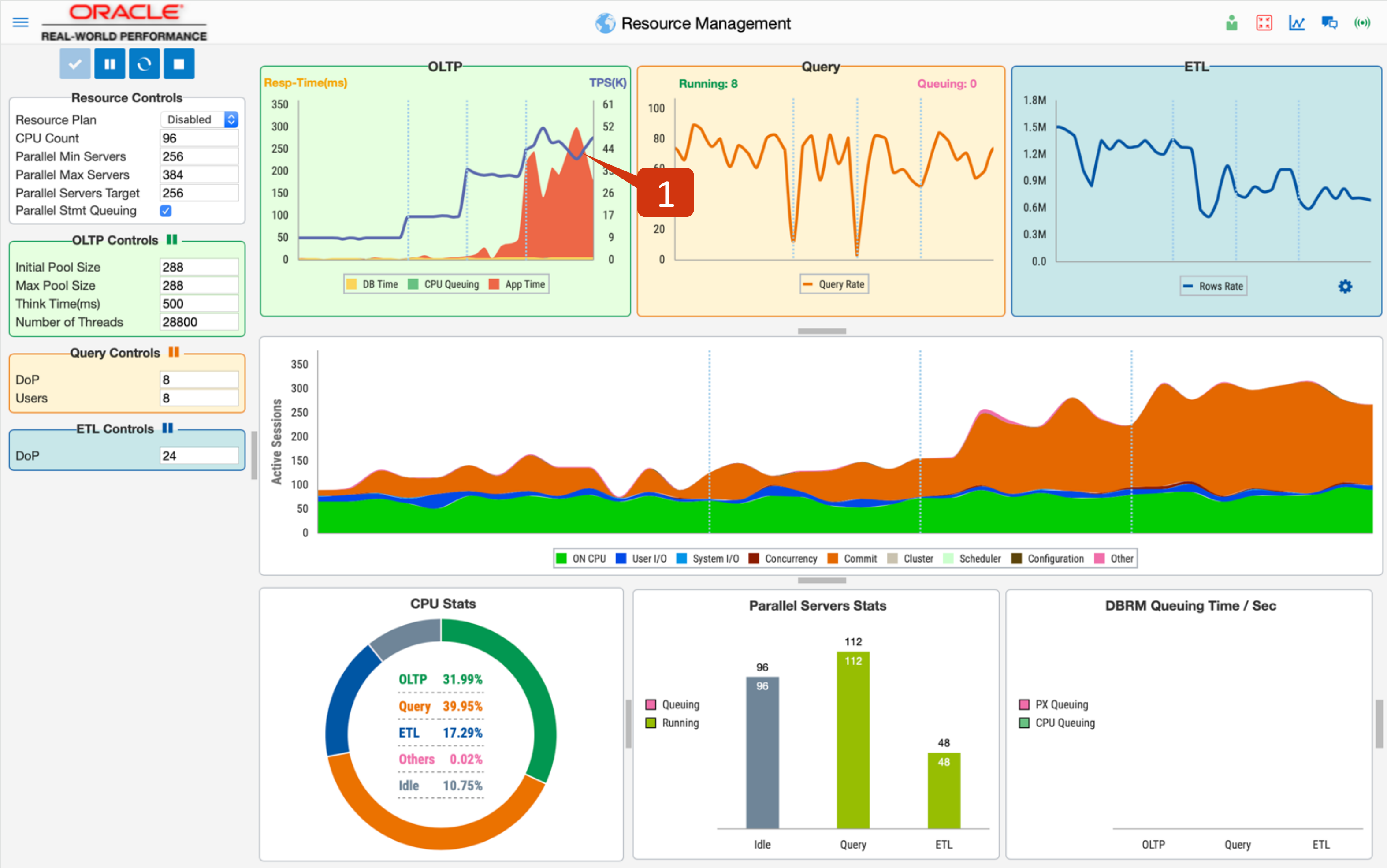
DoP	24

CPU Stats

OLTP	30.23%
Query	36.15%
ETL	18.79%
Others	0.03%
Idle	14.8%

ORACLE
REAL-WORLD PERFORMANCE

1. Significant waiting for connections is now present at all times, degrading overall response times



Resource Management

Resource Controls

Resource Plan	Disabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	500
Number of Threads	28800

Query Controls

DoP	8
Users	8

ETL Controls

DoP	24

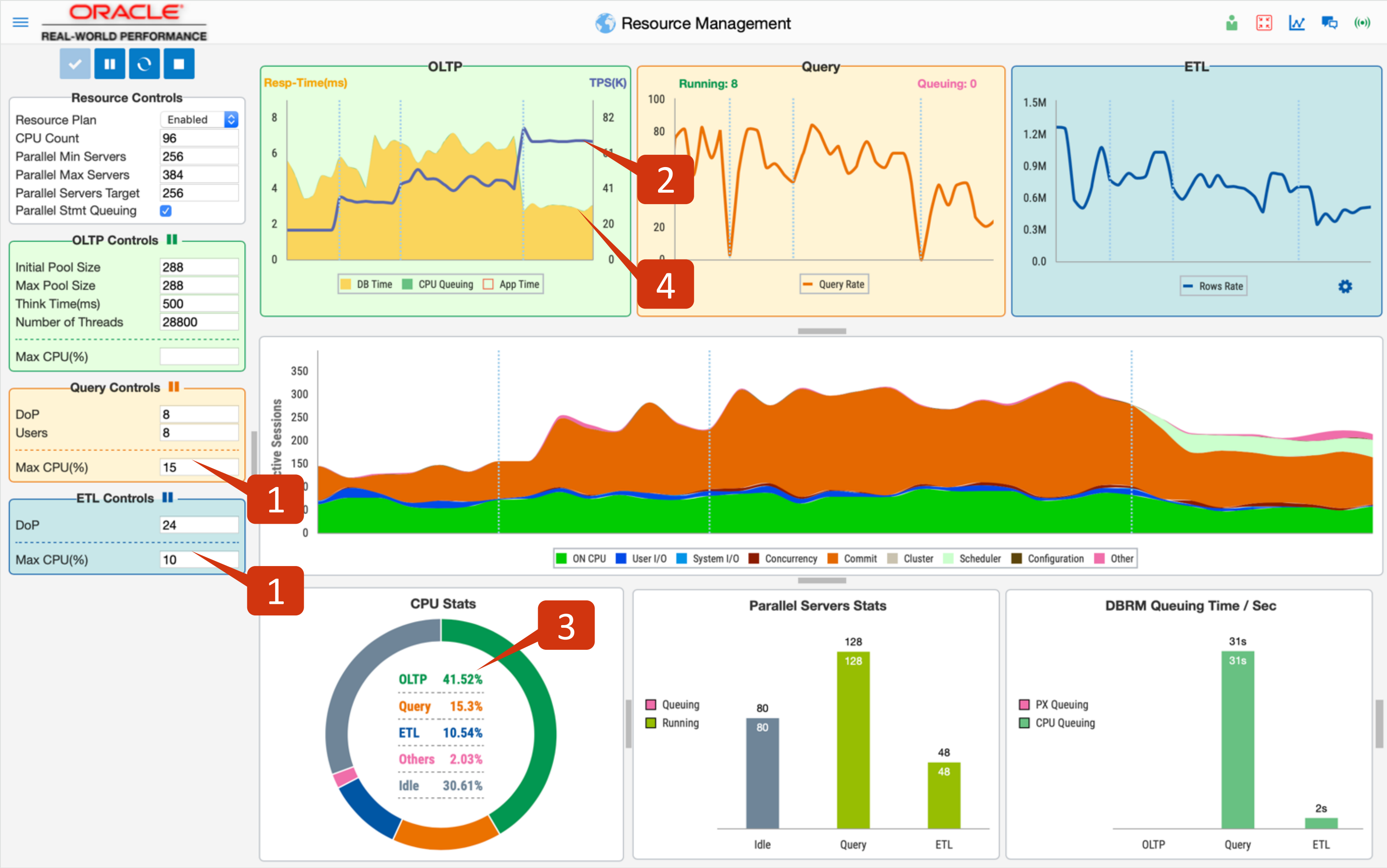
ORACLE
REAL-WORLD PERFORMANCE

Enabling Database Resource Manager

- We can reduce competition for CPU by enabling Database Resource Manager

  Limit the CPU utilization for Query
  Limit the CPU utilization for ETL
  Increase the CPU available for OLTP

1. Enable Database Resource Manager, limiting Query to 15% CPU and limiting ETL to 10% CPU
2. TPS increases
3. OLTP CPU increases
4. Response time in the database improves



Resource Management

Resource Controls

Resource Plan	Enabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	500
Number of Threads	28800
Max CPU(%)	

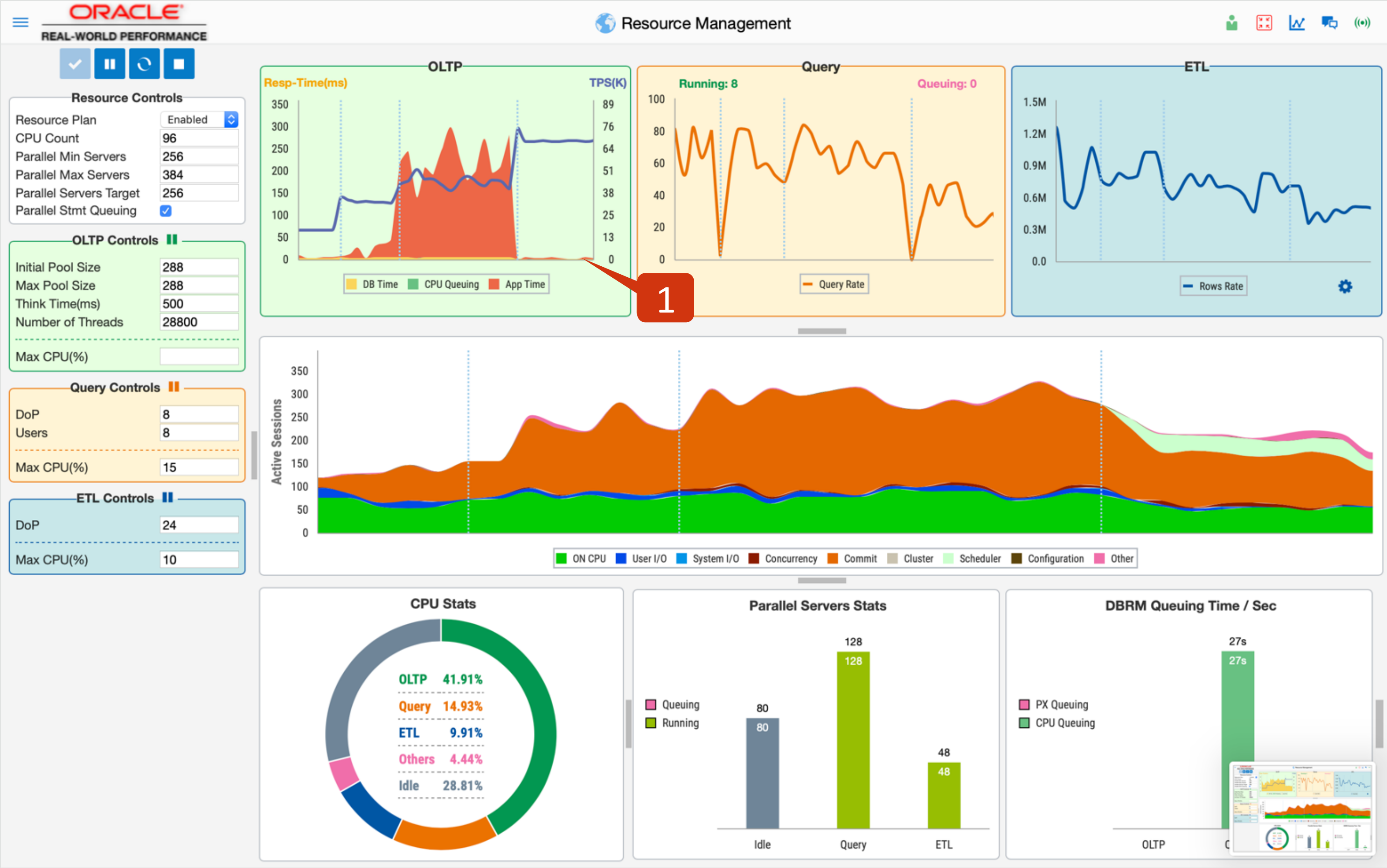
Query Controls

DoP	8
Users	8
Max CPU(%)	15

ETL Controls

DoP	24
Max CPU(%)	10

1. Waiting for connections is significantly reduced, dramatically improving overall response times



ORACLE REAL-WORLD PERFORMANCE

Resource Management

Resource Controls

Resource Plan	Enabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	500
Number of Threads	28800
Max CPU(%)	

Query Controls

DoP	8
Users	8
Max CPU(%)	15

ETL Controls

DoP	24
Max CPU(%)	10

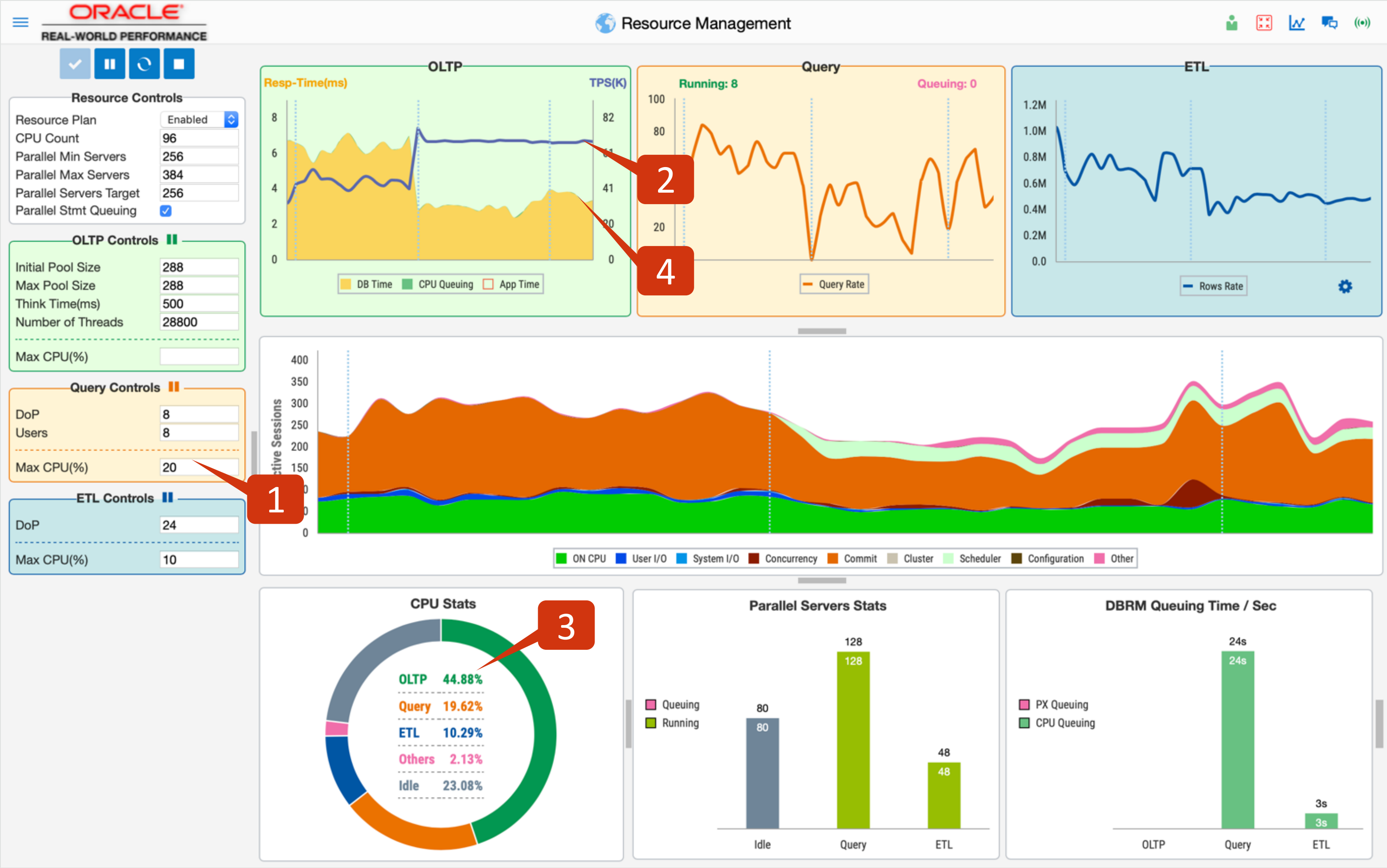
Increasing Competition For CPU

- We can increase competition for CPU by increasing the limits on CPU utilization

  Increasing the limit on CPU utilization for Query increases competition
  Increasing the limit on CPU utilization for ETL increases competition
  Increasing CPU utilization may degrade OLTP performance

ORACLE
REAL-WORLD PERFORMANCE

1. Increase Limit on Query to 20% CPU
2. TPS is unchanged
3. OLTP CPU is about the same
4. Response time in the database increases a little



ORACLE
REAL-WORLD PERFORMANCE

Resource Management

Resource Controls

Resource Plan	Enabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	500
Number of Threads	28800
Max CPU(%)	

Query Controls

DoP	8
Users	8
Max CPU(%)	20

ETL Controls

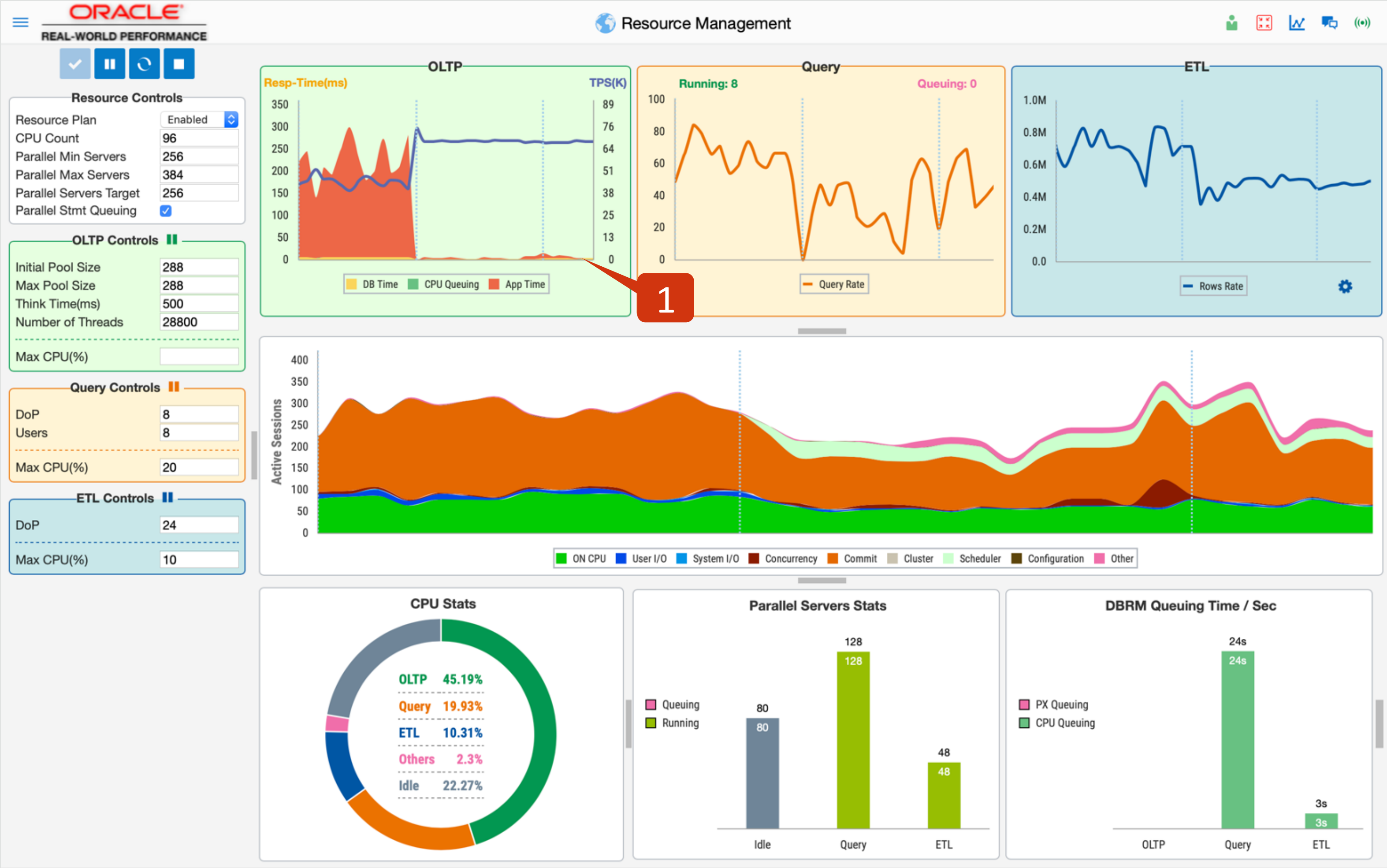
DoP	24
Max CPU(%)	10

Query: Running: 8 Queuing: 0

CPU Stats

OLTP	44.88%
Query	19.62%
ETL	10.29%
Others	2.13%
Idle	23.08%

1. Waiting for connections increases a little



Resource Management

Resource Controls

Resource Plan: Enabled
CPU Count: 96
Parallel Min Servers: 256
Parallel Max Servers: 384
Parallel Servers Target: 256
Parallel Stmt Queuing: ☑

OLTP Controls

Initial Pool Size: 288
Max Pool Size: 288
Think Time(ms): 500
Number of Threads: 28800
Max CPU(%):

Query Controls

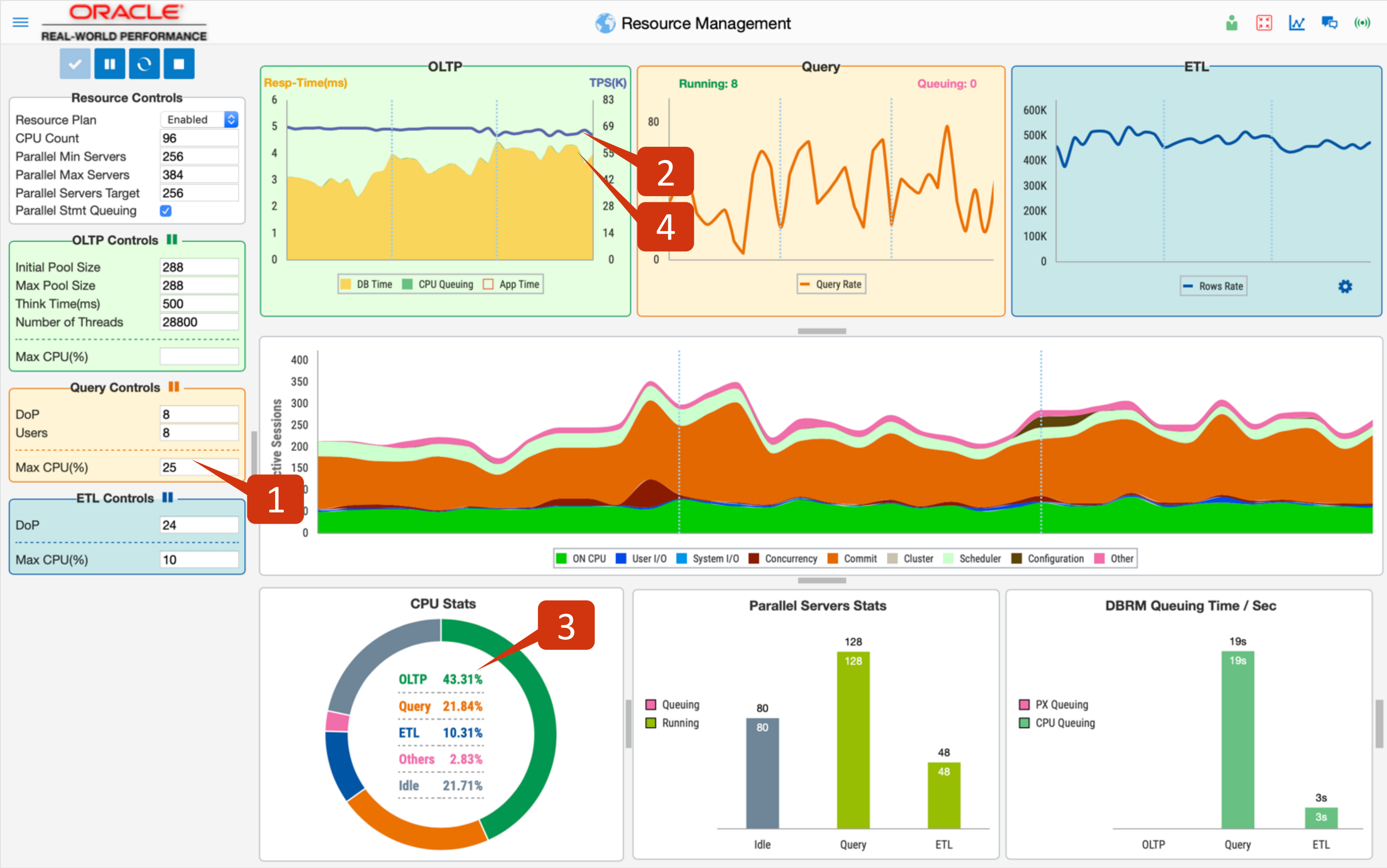
DoP: 8
Users: 8
Max CPU(%): 20

ETL Controls

DoP: 24
Max CPU(%): 10

ORACLE
REAL-WORLD PERFORMANCE

1. Increase Limit on Query to 25% CPU
2. TPS becomes erratic
3. OLTP CPU is about the same
4. Response time in the database increases


Resource Management

Resource Controls

Resource Plan	Enabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	500
Number of Threads	28800
Max CPU(%)	

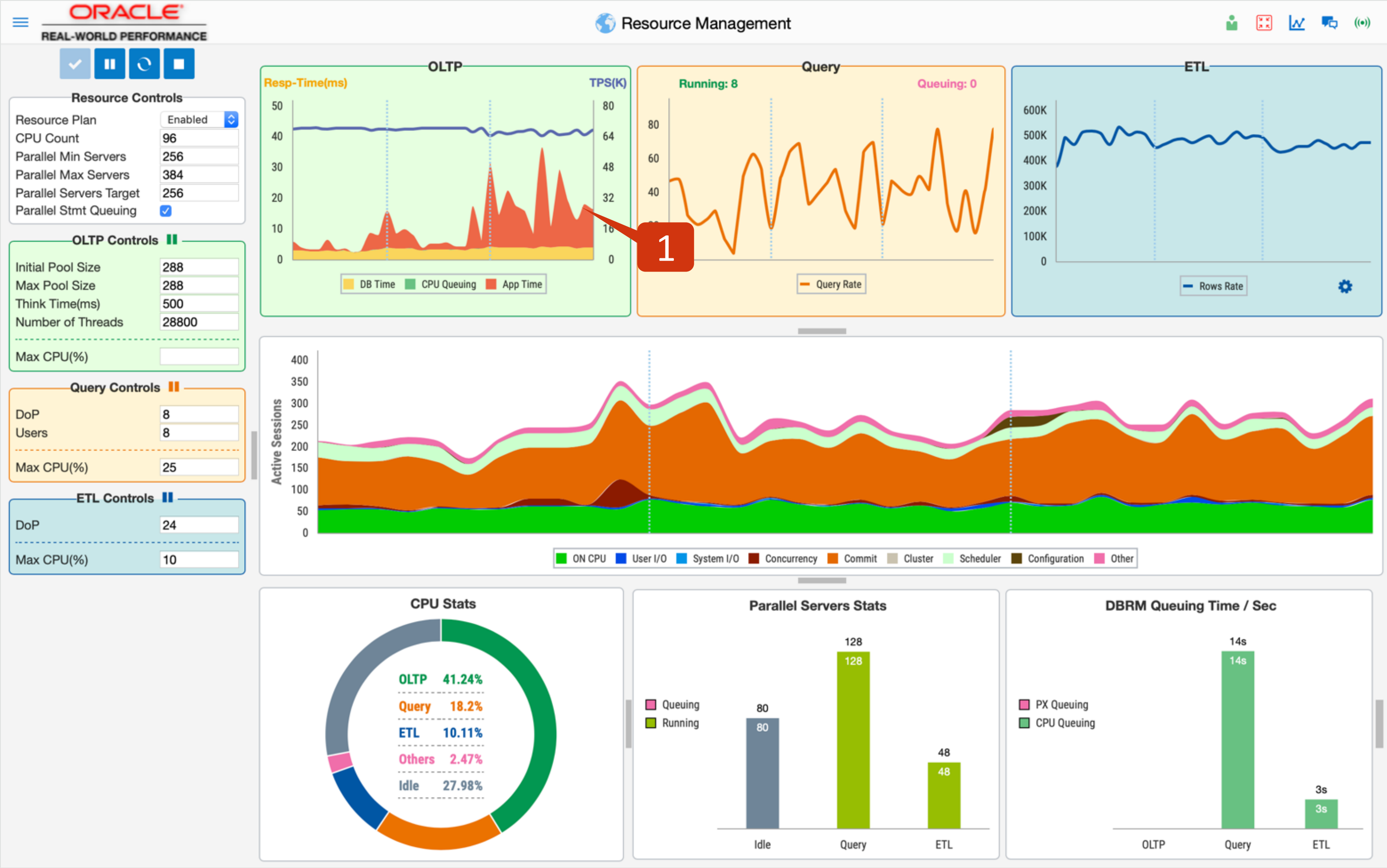
Query Controls

DoP	8
Users	8
Max CPU(%)	25

ETL Controls

DoP	24
Max CPU(%)	10

1. Waiting for connections is now present at all times, degrading overall response times



Resource Management

Resource Controls

Resource Plan: Enabled
CPU Count: 96
Parallel Min Servers: 256
Parallel Max Servers: 384
Parallel Servers Target: 256
Parallel Stmt Queuing: ☑

OLTP Controls

Initial Pool Size: 288
Max Pool Size: 288
Think Time(ms): 500
Number of Threads: 28800
Max CPU(%):

Query Controls

DoP: 8
Users: 8
Max CPU(%): 25

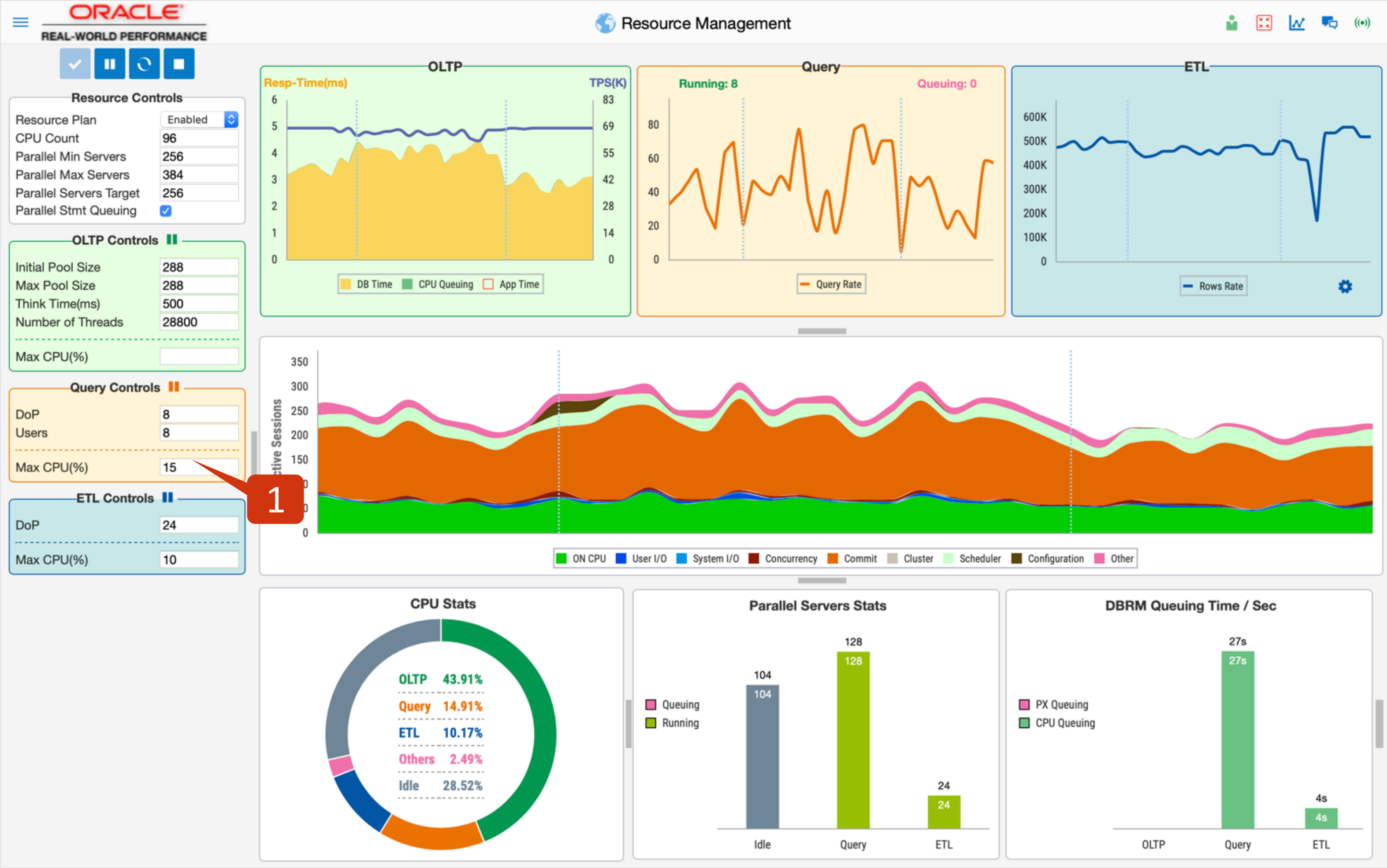
ETL Controls

DoP: 24
Max CPU(%): 10

Decreasing Competition For CPU

- We can decrease competition for CPU by decreasing the limits on CPU utilization
  - Decreasing the limit on CPU utilization for Query decreases competition
  - Decreasing the limit on CPU utilization for ETL decreases competition
  - Decreasing CPU utilization may improve OLTP performance
- We can check for headroom by changing Think Time
  - Decreasing Think Time increases demand
  - DBRM can limit the competition for CPU due to Query and ETL

1. Reset Limit on Query to 15% CPU



Resource Controls

Resource Plan	Enabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	500
Number of Threads	28800
Max CPU(%)	

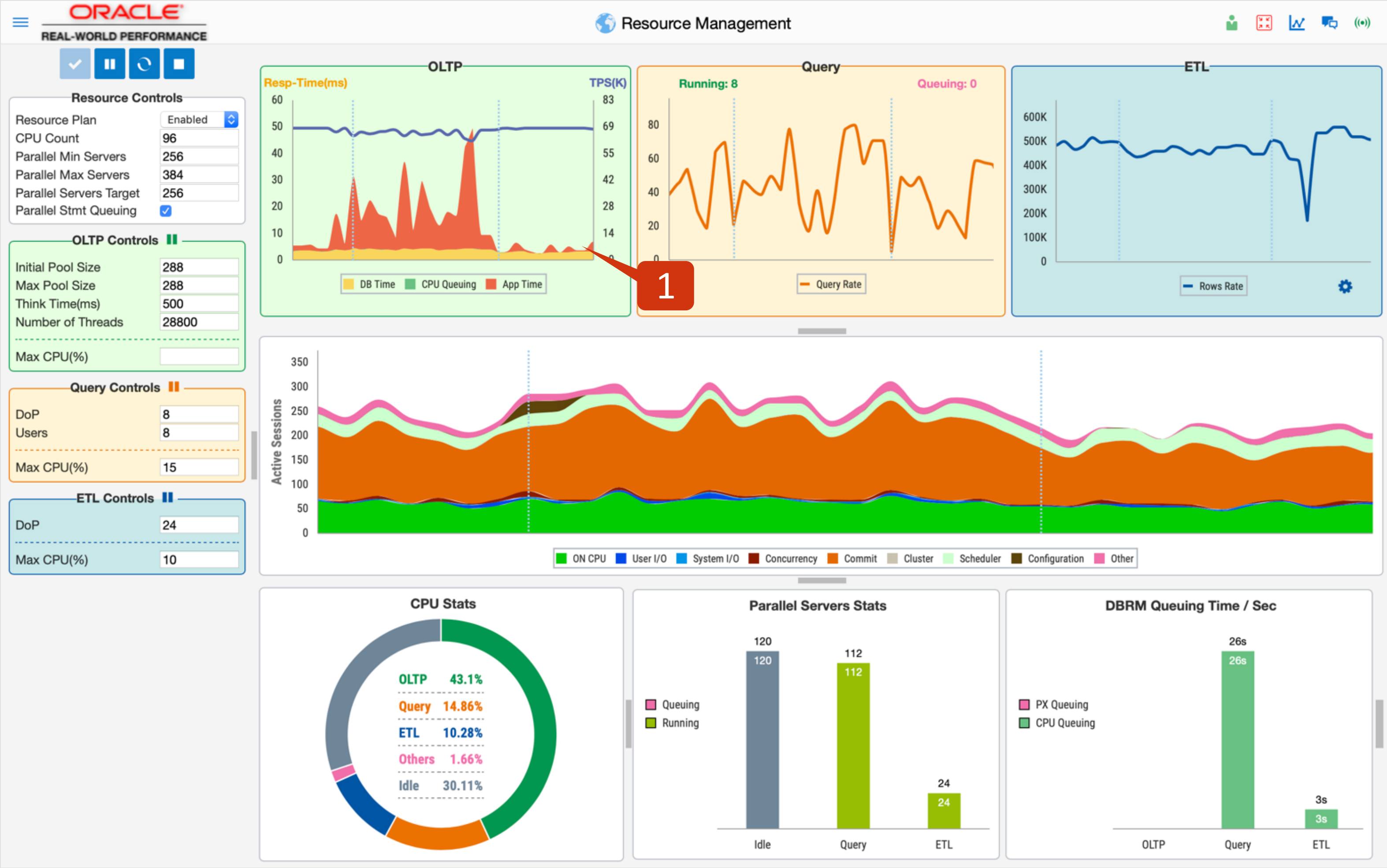
Query Controls

DoP	8
Users	8
Max CPU(%)	15

ETL Controls

DoP	24
Max CPU(%)	10

1. Waiting for connections is significantly reduced, improving overall response times



Resource Management

Resource Controls

Resource Plan	Enabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	500
Number of Threads	28800
Max CPU(%)	

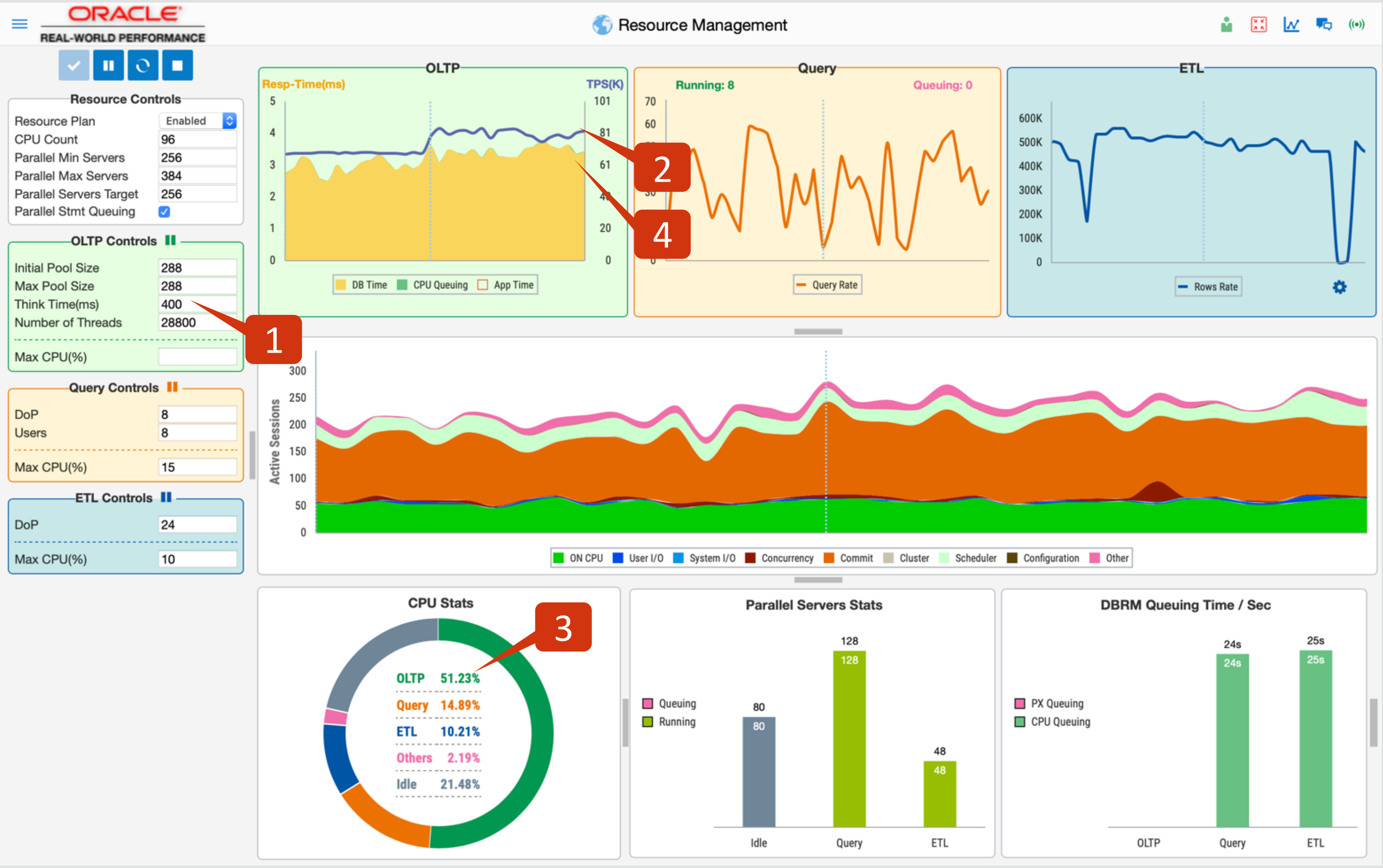
Query Controls

DoP	8
Users	8
Max CPU(%)	15

ETL Controls

DoP	24
Max CPU(%)	10

1. Adjust Think Time to 400 ms to check for headroom
2. TPS increases but becomes erratic
3. OLTP CPU increases
4. Response time in the database also increases a little



Resource Management

Resource Controls

Resource Plan	Enabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	400
Number of Threads	28800
Max CPU(%)	

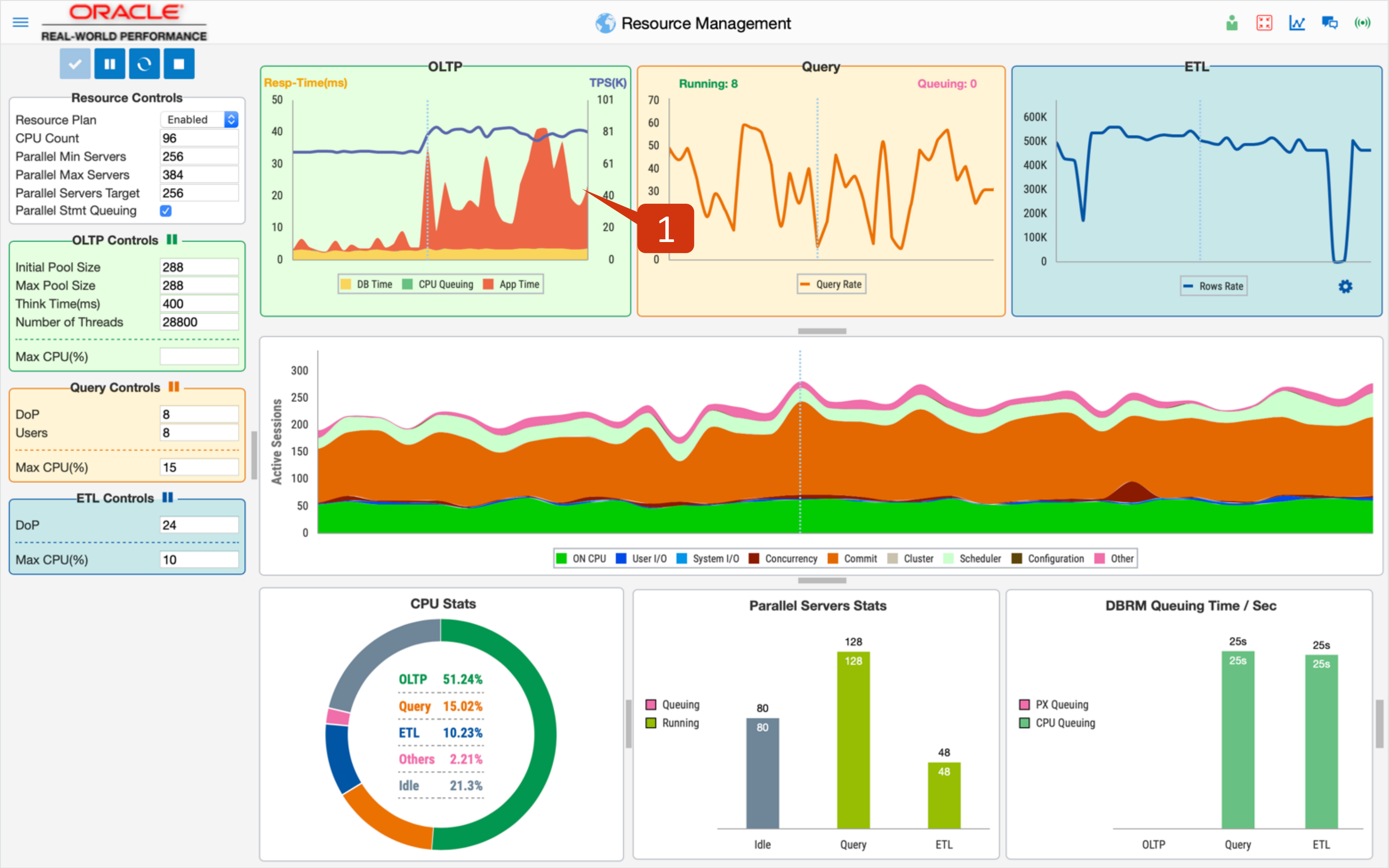
Query Controls

DoP	8
Users	8
Max CPU(%)	15

ETL Controls

DoP	24
Max CPU(%)	10

1. Significant waiting for connections is now present at all times, degrading overall response times. There is little headroom for increased demand from OLTP.



Resource Controls

Resource Plan	Enabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	400
Number of Threads	28800
Max CPU(%)	

Query Controls

DoP	8
Users	8
Max CPU(%)	15

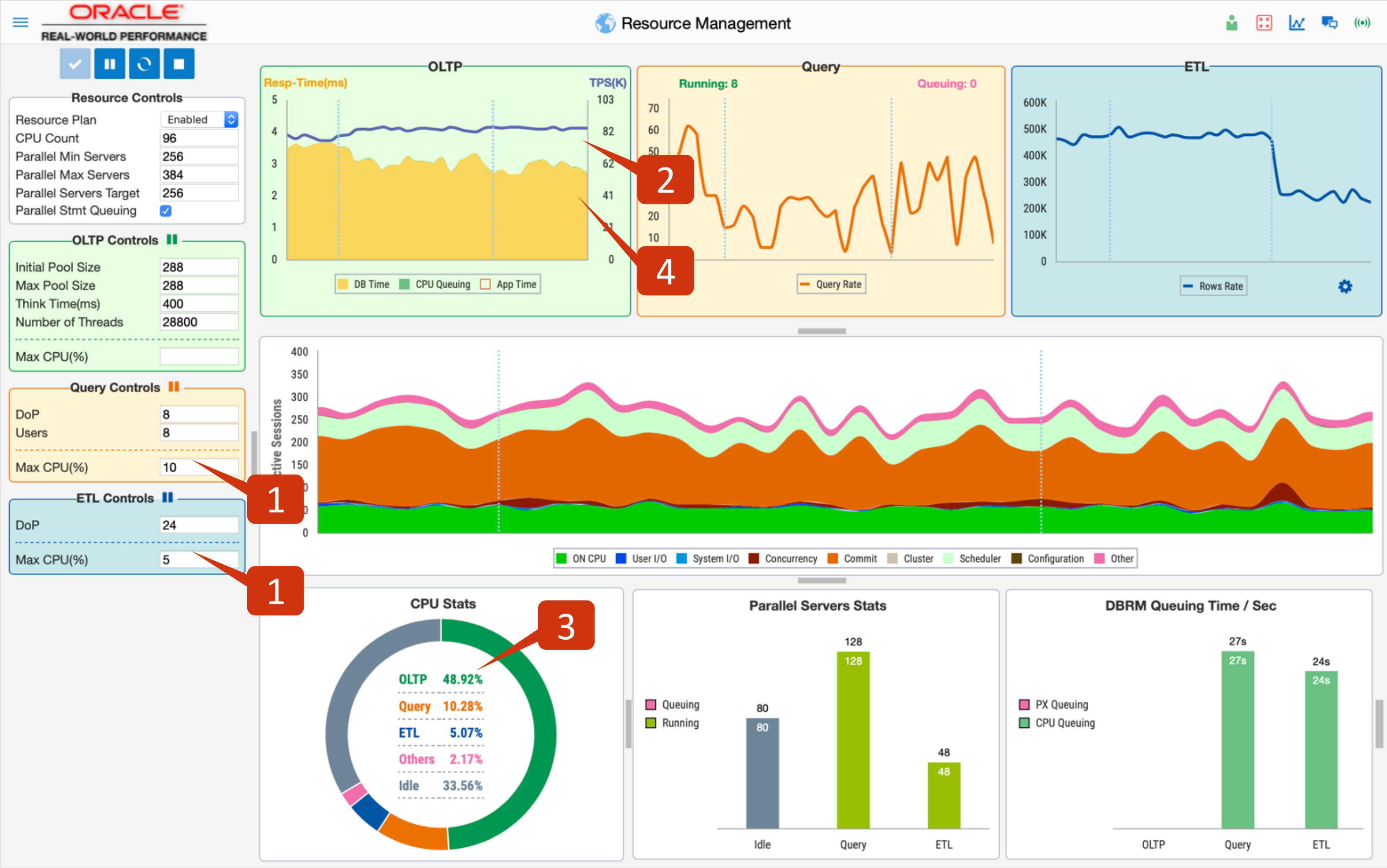
ETL Controls

DoP	24
Max CPU(%)	10

CPU Stats

OLTP	51.24%
Query	15.02%
ETL	10.23%
Others	2.21%
Idle	21.3%

1. Reduce Limit on Query to 10% CPU and Limit on ETL to 5% CPU
2. TPS is consistent
3. OLTP CPU increases a little
4. Response time in the database improves a little



Resource Management

Resource Controls

Resource Plan	Enabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	400
Number of Threads	28800
Max CPU(%)	

Query Controls

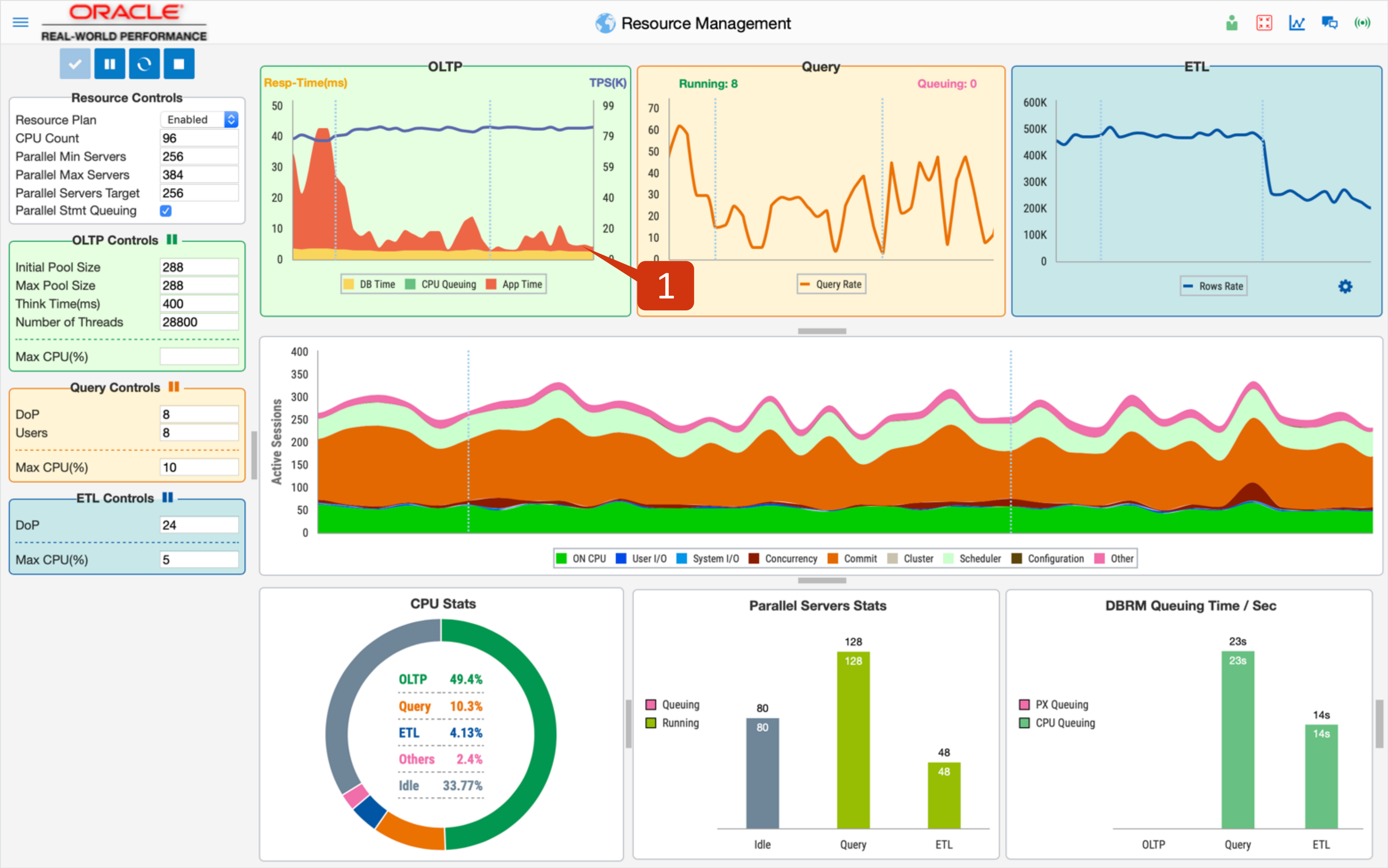
DoP	8
Users	8
Max CPU(%)	10

ETL Controls

DoP	24
Max CPU(%)	5

ORACLE
REAL-WORLD PERFORMANCE

1. Waiting for connections is significantly reduced, dramatically improving overall response times



Resource Management

Resource Controls

Resource Plan	Enabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	400
Number of Threads	28800
Max CPU(%)	

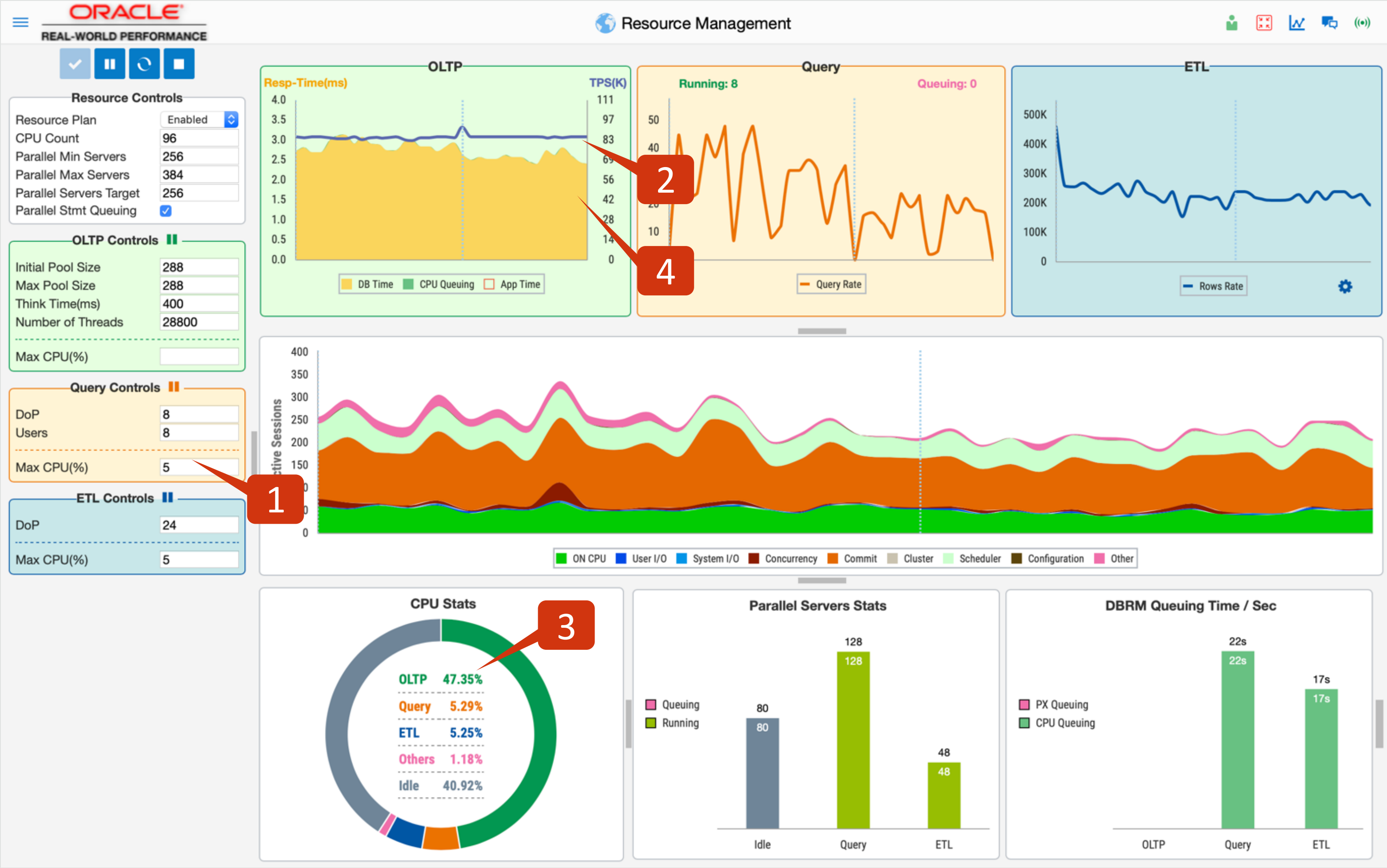
Query Controls

DoP	8
Users	8
Max CPU(%)	10

ETL Controls

DoP	24
Max CPU(%)	5

1. Reduce Limit on Query to 5% CPU
2. TPS is consistent
3. OLTP CPU is about the same
4. Response time in the database improves a little



Resource Management

Resource Controls

Resource Plan	Enabled
CPU Count	96
Parallel Min Servers	256
Parallel Max Servers	384
Parallel Servers Target	256
Parallel Stmt Queuing	☑

OLTP Controls

Initial Pool Size	288
Max Pool Size	288
Think Time(ms)	400
Number of Threads	28800
Max CPU(%)	

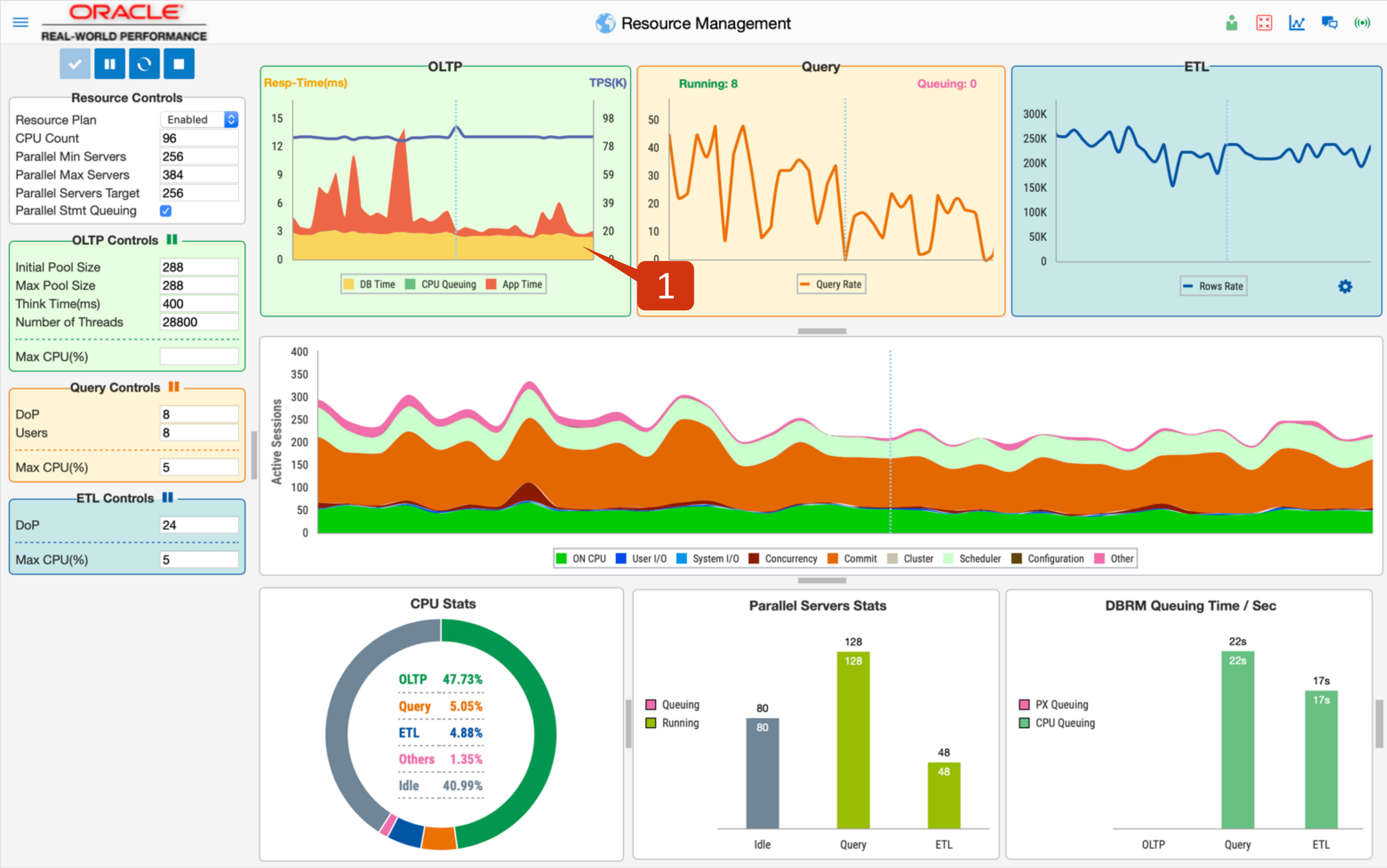
Query Controls

DoP	8
Users	8
Max CPU(%)	5

ETL Controls

DoP	24
Max CPU(%)	5

1. Waiting for connections is reduced, improving overall response times



Resource Management

Resource Controls

Resource Plan: Enabled
CPU Count: 96
Parallel Min Servers: 256
Parallel Max Servers: 384
Parallel Servers Target: 256
Parallel Stmt Queuing: ☑

OLTP Controls

Initial Pool Size: 288
Max Pool Size: 288
Think Time(ms): 400
Number of Threads: 28800
Max CPU(%):

Query Controls

DoP: 8
Users: 8
Max CPU(%): 5

ETL Controls

DoP: 24
Max CPU(%): 5

Agenda

1 Types of Workloads

2 Challenges of Mixed Workloads

3 Ways to Manage Mixed Workloads

4 Demo

5 Lessons Learned

How Do You Determine the Workload CPU Limits?

1. Do not limit the critical OLTP workload
2. Determine the peak OLTP CPU Utilization % by itself
3. Use a simple formula to estimate the CPU available for other workloads

60~70 – ((OLTP CPU %) * 1.3) = CPU for others

Example

- OLTP peak CPU Utilization is 30% by itself
- Lower limit
  60 – ((30* 1.3) = Other workloads can use up to ~20% CPU
- Upper limit
  70 – ((30* 1.3) = Other workloads can use up to ~30% CPU

Lessons Learned

1. Need to limit system CPU utilization to protect OLTP workloads
2. Avoid CPU contention from other workloads
3. Limits give you the highest degree of control

Thank You

Mihajlo Tekic
John Zimmerman

Real-World Performance Team
Oracle Database Development

ORACLE
REAL-WORLD PERFORMANCE