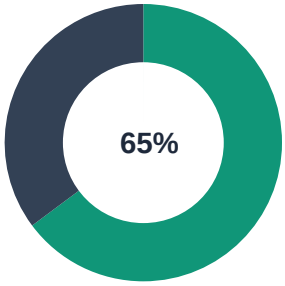
What we found

You can save about $159,690 a year (65%) by moving 2 of your 5 servers to a smaller size. CPU, memory and disk activity still fit.

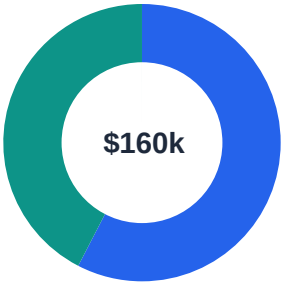
Saving a year	Spend today	Spend after	Servers priced
$159,690	**$246,400**	**$86,720**	**2 of 3**
65% of today's spend	$20,533 a month	$7,226 a month	optimized servers

Today's spend



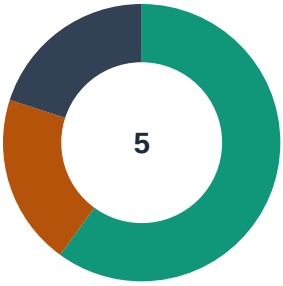
- Saving: $159,690 (65%)
- Still paid: $86,720 (35%)

Where the saving comes from



- Licensing: $91,990 (58%)
- Instance: $67,700 (42%)

Your servers

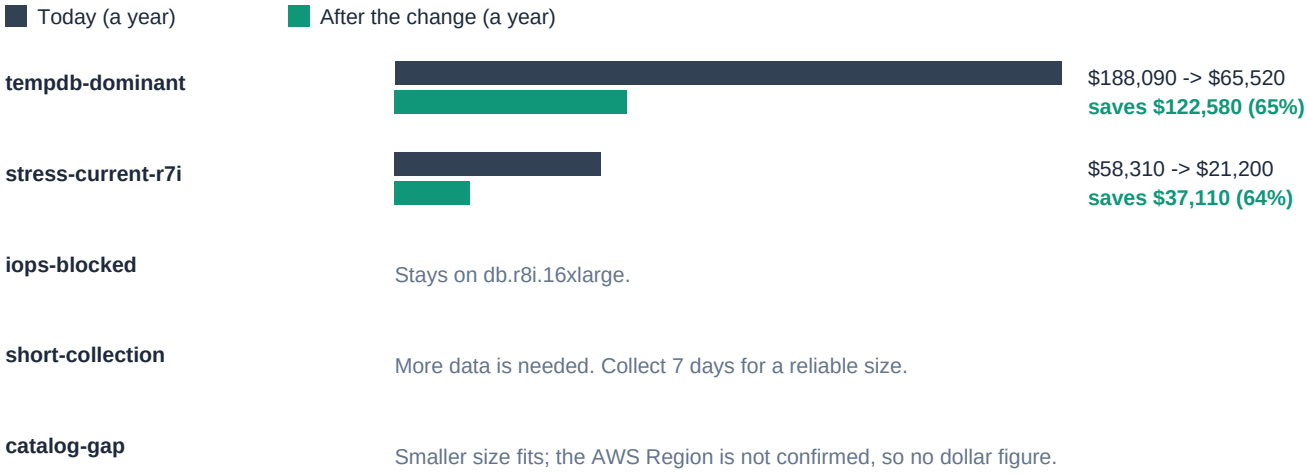


- Optimized: 3
- More data needed: 1
- As is: 1

Why

- The servers that can move use at most 22% of their CPU at busy times (the 95th percentile), so they need fewer vCPU: 72 vCPU today, 24 after. Savio checked memory, disk activity, tempdb and AWS availability too; a separate independent check agrees.
- AWS charges the SQL Server licence for every vCPU, every hour. Fewer vCPU lower that bill by about $91,990 a year.
- Smaller sizes also cost less to run: about $67,700 a year less in instance price.
- 1 server stays as is: a smaller size would not fit their disk activity.
- 1 server needs a longer collection: Savio needs at least 7 days of readings to size a server safely.

Spend today and after, by server



iops-blocked Stays on db.r8i.16xlarge.

short-collection More data is needed. Collect 7 days for a reliable size.

catalog-gap Smaller size fits; the AWS Region is not confirmed, so no dollar figure.

What to do next

- iops-blocked: keep db.r8i.16xlarge; no change needed.
- short-collection: run the collector for at least 7 days and upload again.
- catalog-gap: test the workload on db.r8i.4xlarge with Optimize CPU at 7 vCPU first, then change.
- tempdb-dominant: take a snapshot, then change to db.m8i.8xlarge with Optimize CPU at 11 vCPU in a maintenance window (saves about $122,580 a year).
- stress-current-r7i: take a snapshot, then change to db.m8i.4xlarge with Optimize CPU at 6 vCPU in a maintenance window (saves about $37,110 a year).

About these figures: they show the saving from making this change, at AWS's public on-demand list price in your Region. They may not match your actual AWS spending: discounts, Reserved Instances and Savings Plans lower the instance part (not the SQL Server licensing part), and your workload may change over time. Use them to compare options, not as a forecast of your bill. Assumes the SQL Server license is included in your AWS price (the RDS default). If you bring your own licenses, the licensing part does not apply to your AWS bill. If you have a Reserved Instance for the current size, the instance saving starts when it ends.

Each server's checks and measurements follow.

Business Decision

As is: keep db.r8i.16xlarge

Collection: 7 days, high confidence

Current	Target
db.r8i.16xlarge	db.r8i.16xlarge
The current size stays.	

Optimization shape

Visible vCPU



Lowest memory allowed



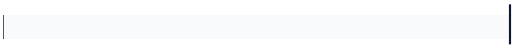
Workload checks

Projected SQL CPU P95



67.6 / 70 %

Projected SQL CPU P99



n/a / 90 %

IOPS P95



50000 / 60000 IOPS

Throughput P95



850 / 1250 MiB/s

Meaningful business signals

- Key workload signals: IOPS fit failure (observed 50000 IOPS; safe capacity 60000 IOPS)
- Top database driver: orders.
- Full technical details are in the CSV and JSON files.
- Edition opportunity: no separate edition migration is included in this PDF.

No saving figure: this server keeps its current size.

Result and next steps

Disk activity is too high for a smaller size. On the next smaller size it would need 50,000 IOPS; that size allows 42,000 IOPS. Based on 7 days of data.

Next steps: No change needed now; If you reduce disk activity on orders, collect again and upload.

Storage

Storage is assessed separately from compute; its saving is not included in the compute figures.

Storage validation required

Current vs assessed storage

Current	Assessed
gp3 | 2048 GiB	**No lower target**

Allocated storage is not reduced.

Provisioning reduction

Provisioned IOPS



Current 64000 IOPS

Target 64000 IOPS

Provisioned throughput



Current 1500 MiB/s

Target 1500 MiB/s

Observed demand and capacity

IOPS P95



Observed 50000 IOPS

IOPS P99



Observed 50000 IOPS

Throughput P95



Observed 850 MiB/s

Throughput P99



Observed 850 MiB/s

Business interpretation

- No lower storage provisioning target was established from the available evidence.
- Latency status: observed (read P95 2 ms, write P95 2 ms, total P99 2 ms).
- Top IOPS database driver: orders.

Conditions before action

- AWS does not offer db.r8i.16xlarge for SQL Server Standard Edition ve 16.00.4125.3 in us-east-1, so the storage limits of the current class cou

Planning guidance only. No automated RDS storage modification is performed.

Business Decision

More data needed for a reliable size

Collection: 6 hours, preliminary confidence

Current	Target
db.r8i.16xlarge	db.m8i.8xlarge

Early sign only; more data needed.

Optimization shape

Visible vCPU



Lowest memory allowed vs target



Workload checks

Projected SQL CPU P95



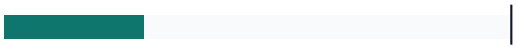
64 / 70 %

Projected SQL CPU P99



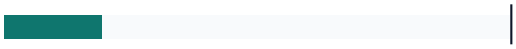
64 / 90 %

IOPS P95



11000 / 40000 IOPS

Throughput P95



240 / 1250 MiB/s

Meaningful business signals

- Key workload signals: insufficient evidence window (observed n/a; safe capacity n/a)
- Top database driver: orders.
- Full technical details are in the CSV and JSON files.
- Edition opportunity: no separate edition migration is included in this PDF.

Saving in dollars not shown: AWS publishes no price for db.r8i.16xlarge with SQL Server Standard Edition in this Regio

Possible: up to 69% fewer licensed vCPU (32 -> 10), confirmed after a longer collection

Result and next steps

Only 6 hours were collected. The size shown is not reliable: more data is needed. Collect 7 days for a reliable size.

Next steps: Run the collector for 7 days; Upload the new ZIP.

Storage

Storage is assessed separately from compute; its saving is not included in the compute figures.

Storage validation required

Current vs assessed storage

Current	Assessed
gp3 | 2048 GiB	**No lower target**

Allocated storage is not reduced.

Provisioning reduction

Provisioned IOPS



Current 32000 IOPS

Target 32000 IOPS

Provisioned throughput



Current 1000 MiB/s

Target 1000 MiB/s

Observed demand and capacity

IOPS P95



Observed 11000 IOPS

IOPS P99



Observed 11000 IOPS

Throughput P95



Observed 240 MiB/s

Throughput P99



Observed 240 MiB/s

Business interpretation

- No lower storage provisioning target was established from the available evidence.
- Latency status: observed (read P95 2 ms, write P95 2 ms, total P99 2 ms).
- Top IOPS database driver: orders.

Conditions before action

- Storage evidence window is not eligible: 6 hours collected; at least 48 h
  needed for a reliable size, ideally 7 days. No gaps in the readings. Mak

Planning guidance only. No automated RDS storage modification is performed.

Business Decision

Optimized after a check: db.r8i.4xlarge with 7 vCPU

Collection: 7 days, high confidence

Current	Target
db.r8i.12xlarge	**db.r8i.4xlarge**

A smaller size fits after one check.

Optimization shape

Visible vCPU



Lowest memory allowed vs target



Workload checks

Projected SQL CPU P95



61.7 / 70 %

Projected SQL CPU P99



61.7 / 90 %

IOPS P95



12000 / 20000 IOPS

Throughput P95



300 / 625 MiB/s

Meaningful business signals

- Key workload signals: CPU gate passed (observed 61.7 %; safe capacity 70 %); memory gate passed (observed 96 GB; safe capacity 128 GB); IOPS gate passed (observed 12000
- Top database driver: orders.
- Full technical details are in the CSV and JSON files.
- Edition opportunity: no separate edition migration is included in this PDF.

Saving in dollars not shown: Savio could not tell your AWS Region from the server name, and prices differ by Region.

71% fewer licensed vCPU (24 -> 7)

Result and next steps

Your server can likely move from db.r8i.12xlarge (24 vCPU) to db.r8i.4xlarge with Optimize CPU set to 7 vCPU. Check this first: Savio could not tell your AWS Region from the server name catalog-gap.rds.amazonaws.com, so it checked the sizes AWS sells in US East (N. Virginia). Use

Next steps: Use the RDS endpoint as ServerName in SavioServers.csv and upload again; Then, in a maintenance window, change the instance to db.r8i.4xlarge with Optimize CPU at 7 vCPU.

Storage

Storage is assessed separately from compute; its saving is not included in the compute figures.

Storage validation required

Current vs assessed storage

Current	Assessed
unknown | 1478 GiB	No lower target

Allocated storage is not reduced.

Provisioning reduction

Provisioned IOPS

Current n/a IOPS

Target n/a IOPS

Provisioned throughput

Current n/a MiB/s

Target n/a MiB/s

Observed demand and capacity

IOPS P95



Observed 12000 IOPS

IOPS P99



Observed 12000 IOPS

Throughput P95



Observed 300 MiB/s

Throughput P99



Observed 300 MiB/s

Business interpretation

- No lower storage provisioning target was established from the available evidence.
- Latency status: observed (read P95 2 ms, write P95 2 ms, total P99 2 ms).
- Top IOPS database driver: orders.

Conditions before action

- Missing current storage facts: storage type, provisioned IOPS, Multi-AZ

Planning guidance only. No automated RDS storage modification is performed.

Business Decision

Optimized: db.m8i.8xlarge with 11 vCPU

Collection: 7 days, high confidence

Current	Target
db.r8i.16xlarge	**db.m8i.8xlarge**
A smaller size fits every check.	

Optimization shape

Visible vCPU



Lowest memory allowed vs target



Workload checks

Projected SQL CPU P95



64 / 70 %

Projected SQL CPU P99



64 / 90 %

IOPS P95



26000 / 40000 IOPS

Throughput P95



740 / 1250 MiB/s

Meaningful business signals

- Key workload signals: SQL Server edition constraint (observed n/a; safe capacity n/a)
- Top database driver: tempdb.
- Full technical details are in the CSV and JSON files.
- Edition opportunity: Standard Edition is blocked and should stay separate from compute scaling.

Estimated saving: about $122,580 a year (65%)

SQL Server licensing: $77,450/year (66% lower) - Instance: $45,130/year (64% lower)
Now db.r8i.16xlarge $15,675/month -> db.m8i.8xlarge $5,460/month

Result and next steps

Your server can move from db.r8i.16xlarge (32 vCPU) to db.m8i.8xlarge with Optimize CPU set to 11 vCPU. CPU, memory and disk activity all fit, based on 7 days of data.

Next steps: Take an RDS snapshot; In a maintenance window, change the instance to db.m8i.8xlarge with Optimize CPU at 11 vCPU; Watch CPU and disk activity for the first busy days.

Storage

Storage is assessed separately from compute; its saving is not included in the compute figures.

Storage optimized

Current vs assessed storage

Current	Assessed
gp3 | 2048 GiB	**gp3 | 2048 GiB**

Allocated storage is not reduced.

Provisioning reduction

Provisioned IOPS



Current 64000 IOPS

Target 37143 IOPS

Provisioned throughput



Current 2000 MiB/s

Target 1058 MiB/s

Observed demand and capacity

IOPS P95



26000 / 26000.1 IOPS

IOPS P99



26000 / 33428.7 IOPS

Throughput P95



740 / 740.6 MiB/s

Throughput P99



740 / 952.2 MiB/s

Business interpretation

- IOPS provisioning reduction: 42%.
- Throughput provisioning reduction: 47.1%.
- Technical storage provisioning is optimized by the assessed target.
- Latency status: observed (read P95 2 ms, write P95 2 ms, total P99 2 ms).
- Top IOPS database driver: tempdb.

Conditions before action

- All approved storage capacity and evidence checks passed.

Planning guidance only. No automated RDS storage modification is performed.

Business Decision

Optimized: db.m8i.4xlarge with 6 vCPU

Collection: 7 days, high confidence

Current	Target
db.r7i.8xlarge	**db.m8i.4xlarge**

A smaller size fits every check.

Optimization shape

Visible vCPU



Lowest memory allowed vs target



Workload checks

Projected SQL CPU P95



58.7 / 70 %

Projected SQL CPU P99



58.7 / 90 %

IOPS P95



13000 / 20000 IOPS

Throughput P95



320 / 625 MiB/s

Meaningful business signals

- Key workload signals: CPU gate passed (observed 58.7 %; safe capacity 70 %); memory gate passed (observed 64 GB; safe capacity 64 GB); IOPS gate passed (observed 13000
- Top database driver: orders.
- Full technical details are in the CSV and JSON files.
- Edition opportunity: no separate edition migration is included in this PDF.

Estimated saving: about $37,110 a year (64%)

SQL Server licensing: $14,540/year (63% lower) - Instance: $22,570/year (64% lower)
Now db.r7i.8xlarge $4,859/month -> db.m8i.4xlarge $1,767/month

Result and next steps

Your server can move from db.r7i.8xlarge (16 vCPU) to db.m8i.4xlarge with Optimize CPU set to 6 vCPU. CPU, memory and disk activity all fit, based on 7 days of data.
Option 1, stay on r7i: db.r7i.4xlarge (6 vCPU)
Option 2, move to the newest family: db.m8i.4xlarge (6 vCPU)

Next steps: Take an RDS snapshot; In a maintenance window, change the instance to db.m8i.4xlarge with Optimize CPU at 6 vCPU; Watch CPU and disk activity for the first busy days.

Storage

Storage is assessed separately from compute; its saving is not included in the compute figures.

Storage optimized

Current vs assessed storage

Current	Assessed
gp3 | 2048 GiB	**gp3 | 2048 GiB**

Allocated storage is not reduced.

Provisioning reduction

Provisioned IOPS



Current 32000 IOPS

Target 18572 IOPS

Provisioned throughput



Current 1000 MiB/s

Target 458 MiB/s

Observed demand and capacity

IOPS P95



13000 / 13000.4 IOPS

IOPS P99



13000 / 16714.8 IOPS

Throughput P95



320 / 320.6 MiB/s

Throughput P99



320 / 412.2 MiB/s

Business interpretation

- IOPS provisioning reduction: 42%.
- Throughput provisioning reduction: 54.2%.
- Technical storage provisioning is optimized by the assessed target.
- Latency status: observed (read P95 2 ms, write P95 2 ms, total P99 2 ms).
- Top IOPS database driver: orders.

Conditions before action

- All approved storage capacity and evidence checks passed.

Planning guidance only. No automated RDS storage modification is performed.