



Institut de Recherche en Informatique de Toulouse  
CNRS - Toulouse INP - UT - UT Capitole - UT2



# Sufficiency in Cloud Distributed Datacenters

**PhD** as a part of **PEPR Cloud** project (2024-2030)

**Supervised by:** Patricia Stolf, Jean-Marc Pierson, Laurent Lefèvre




Institut de Recherche en Informatique de Toulouse  
CNRS - Toulouse INP - UT - UT Capitole - UT2

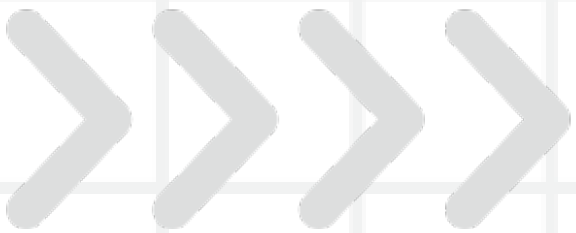


# Sufficiency in Cloud : Towards the Design of an Energy Aware Resource Recommendation System



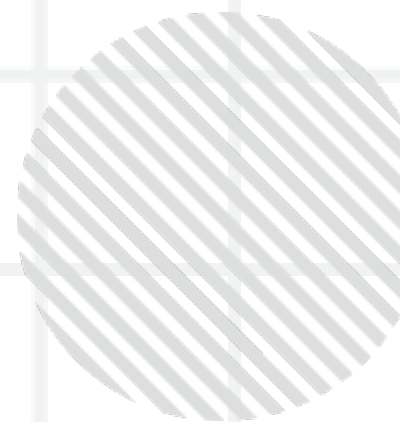
# TABLE OF CONTENTS

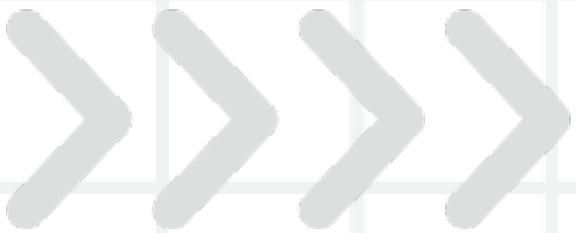
1. **Sufficiency** in Cloud and **User Behavior**
  2. Cluster **Recommendation** System Design
  3. comparec: **Compare** and Recommend
  4. **Evaluation** of Recommendations
  5. Conclusion
- 



# QUESTION

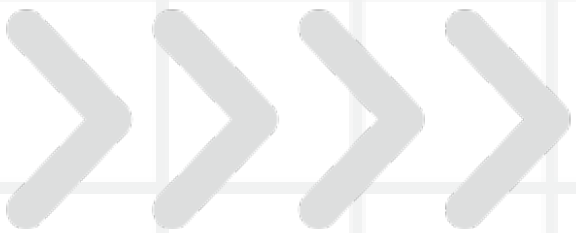
What is **sufficiency in cloud** ?





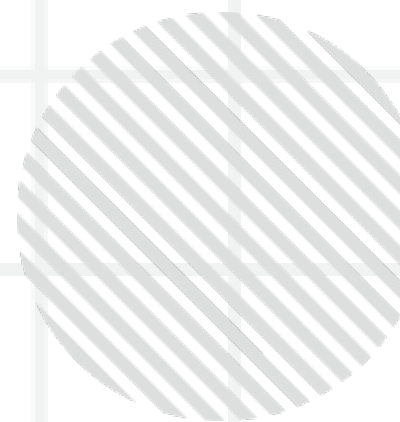
# SUFFICIENCY

1. Focusing on **essentials needs** and **avoiding excess**
2. **Right-sizing** resources
3. Minimizing **near-idle time**
4. Changing user **habits** and **behavior** to be more energy-aware

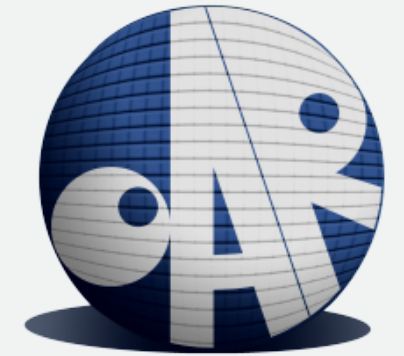


# QUESTION

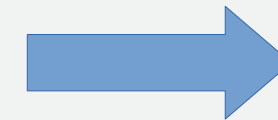
What exactly is **user behavior** ?



# USER BEHAVIOR EXAMPLE

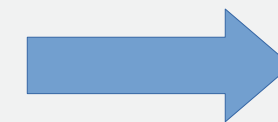


We want to reserve **1 node** from **cluster A**



We are given **1 node** from **cluster A** cluster if it's available. Otherwise, we wait.

We want to reserve **1 node**, but we don't mind which cluster

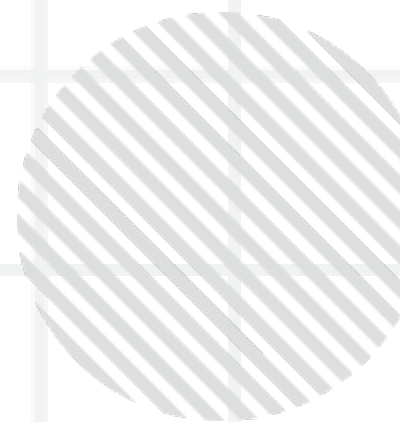


We are given **1 node** from **any cluster** that is available – might not be the best for the job

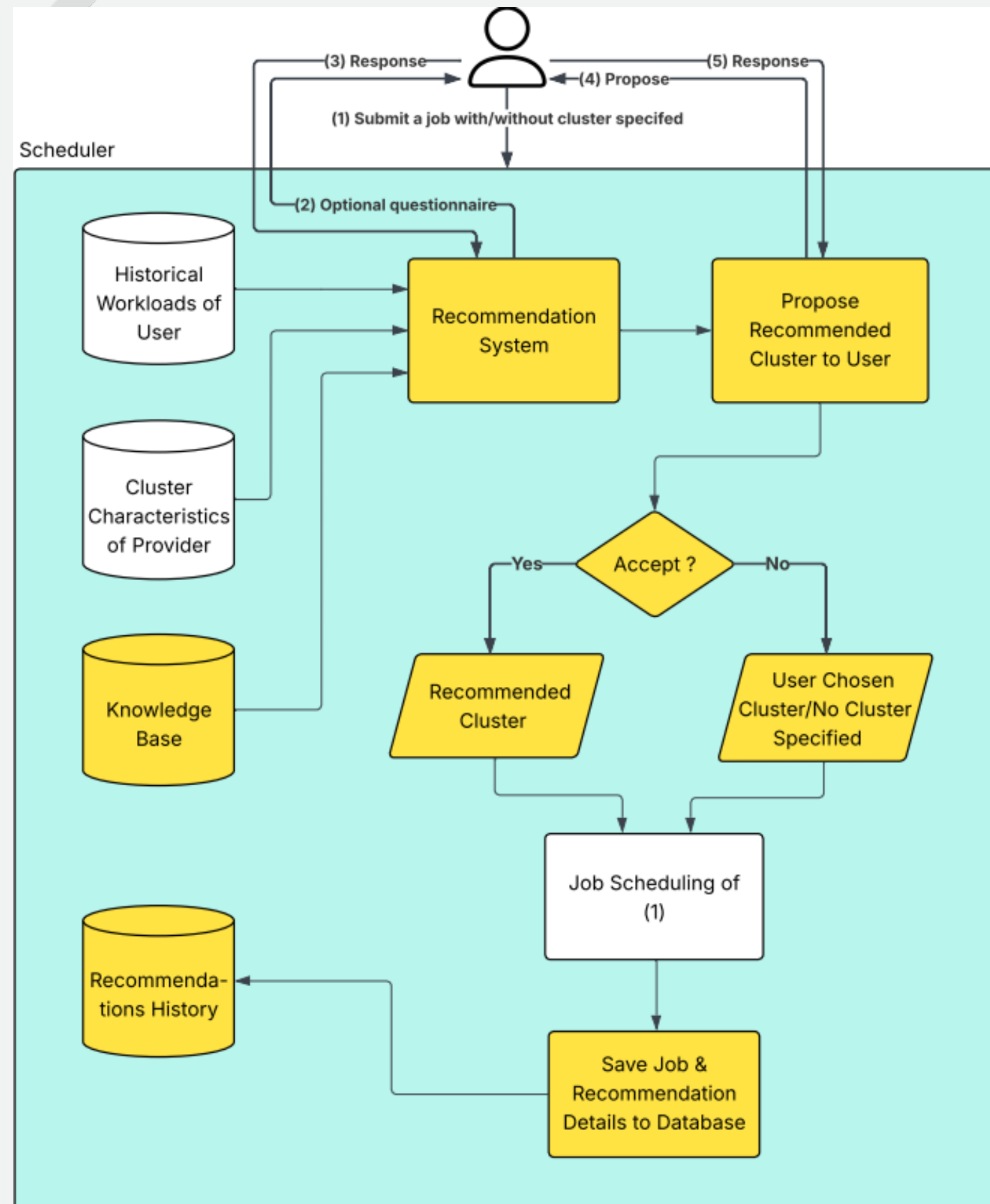


# QUESTION

How can we **assist** users in changing their behavior to be more **energy aware**?



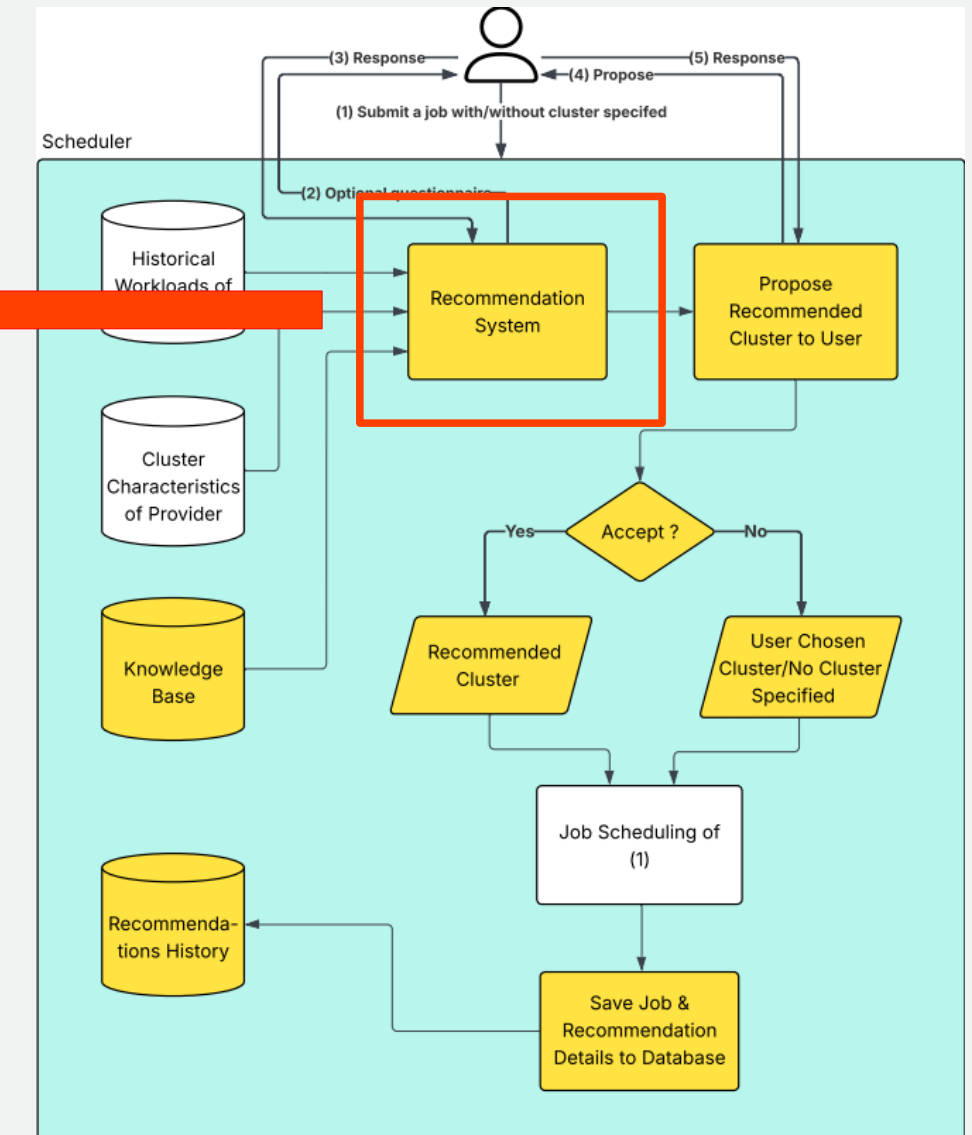
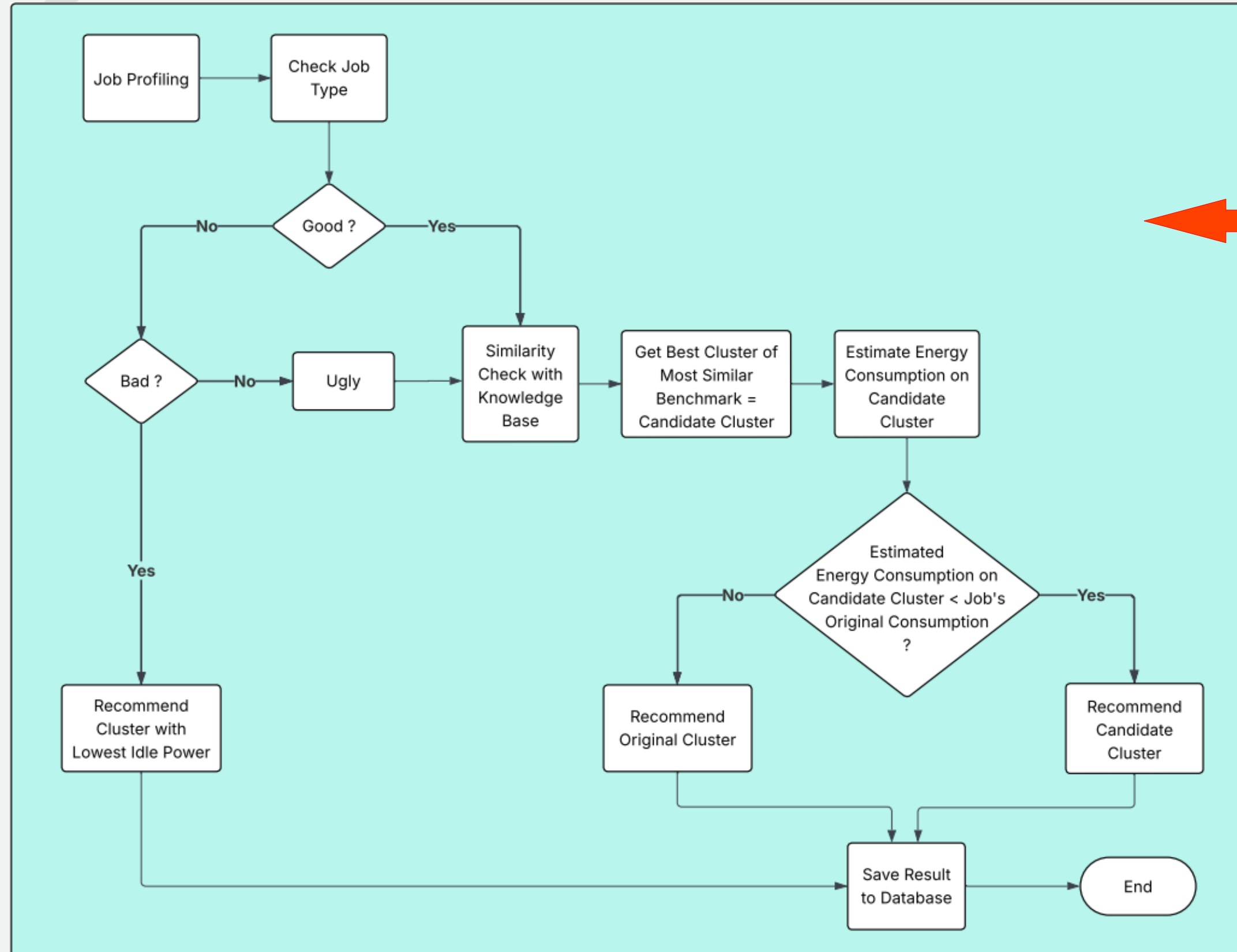
# Cluster Recommendation System Design



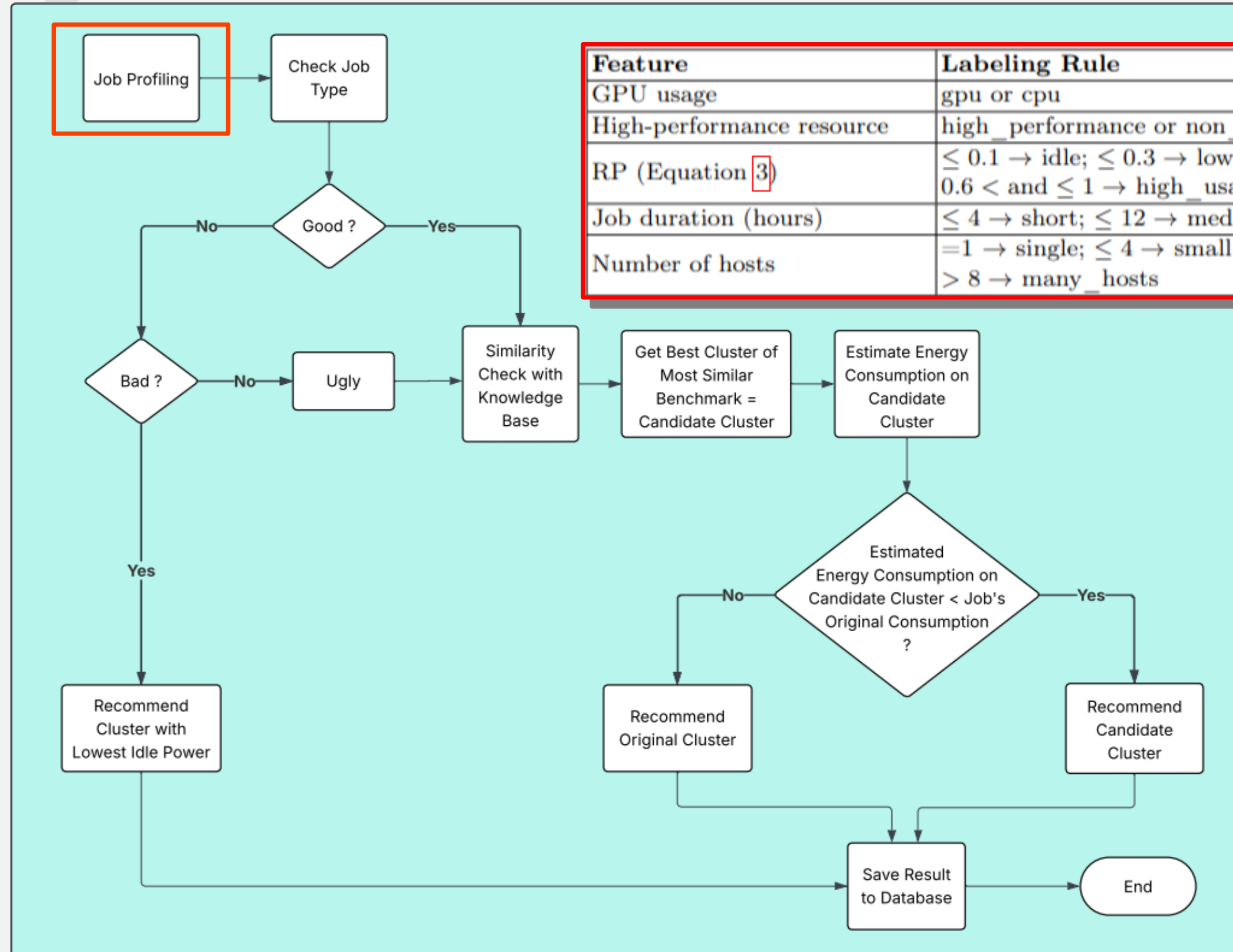
Make a recommendation using 4 main sources of information:

- (1) the user's **historical workloads**;
- (2) a **knowledge base** – set of benchmarks executed on all clusters;
- (3) **cluster characteristics** – idle/max power of each cluster, hardware specs;
- (4) **optional questionnaire** – *whether job is CPU/memory intensive etc.*

# comparec: Compare and Recommend

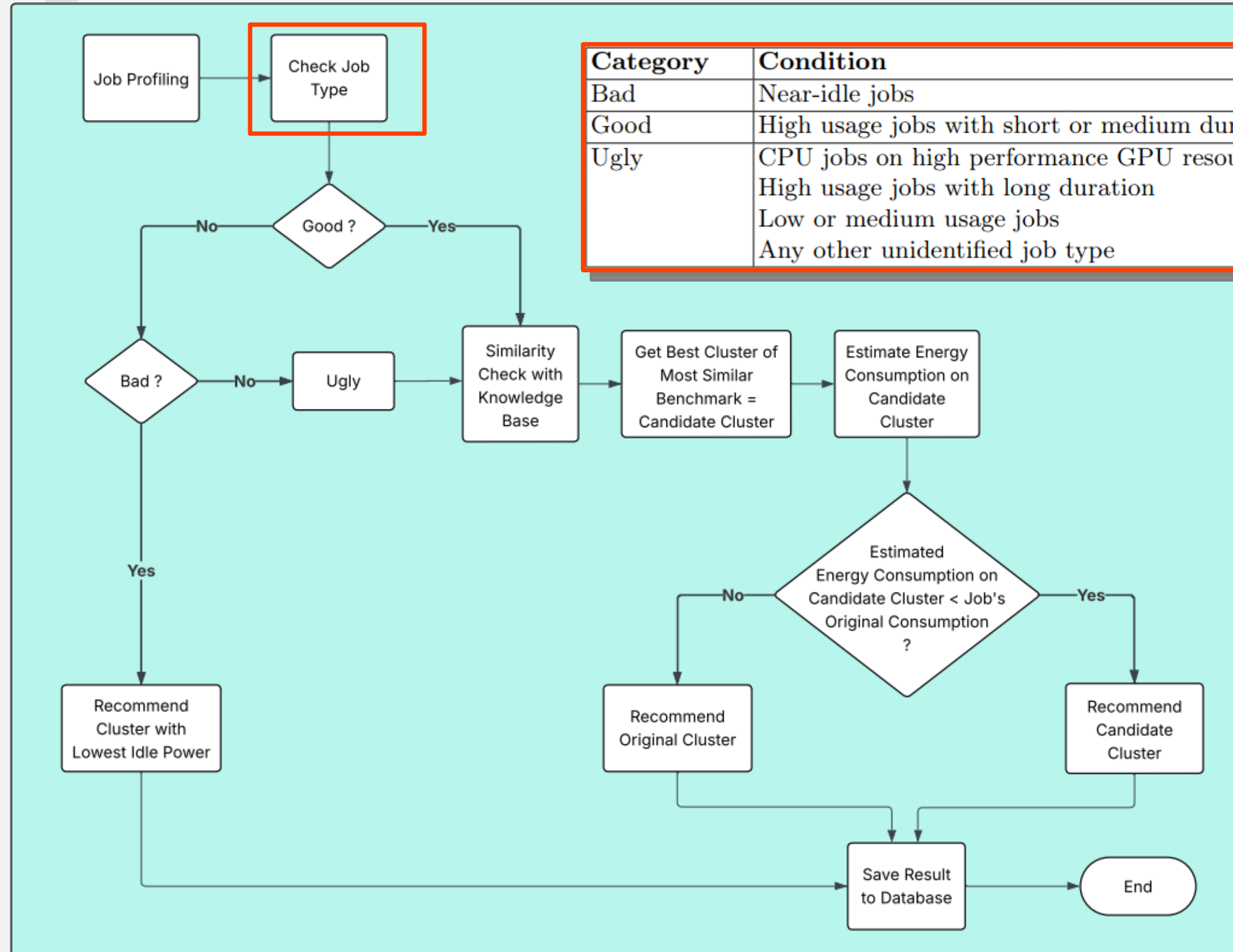


# comparec: Compare and Recommend



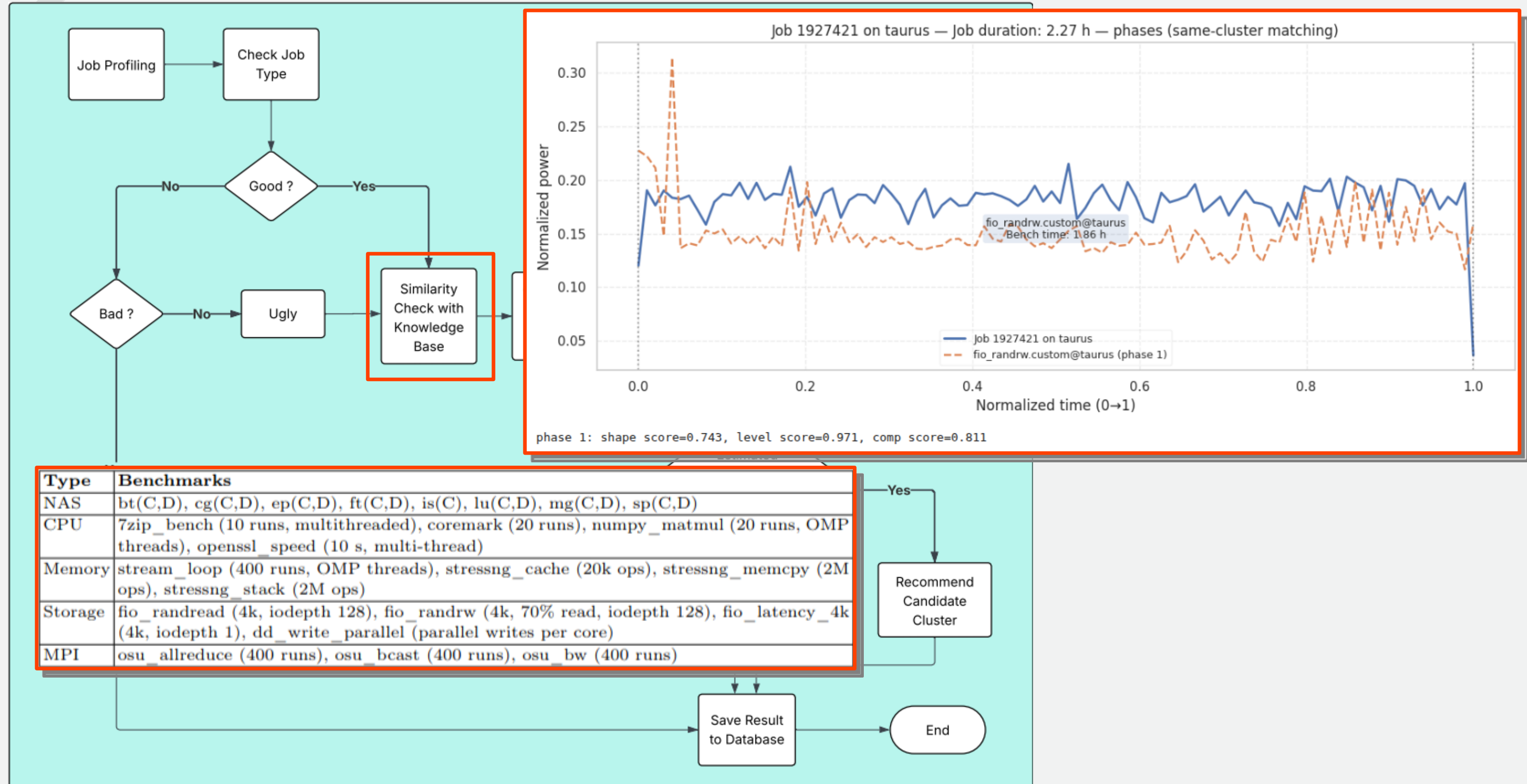
Feature	Labeling Rule
GPU usage	gpu or cpu
High-performance resource	high_performance or non_high_performance
RP (Equation 3)	$\leq 0.1 \rightarrow$ idle; $\leq 0.3 \rightarrow$ low; $\leq 0.6 \rightarrow$ medium; $0.6 < \text{and } \leq 1 \rightarrow$ high_usage
Job duration (hours)	$\leq 4 \rightarrow$ short; $\leq 12 \rightarrow$ medium; $> 12 \rightarrow$ long_duration
Number of hosts	$=1 \rightarrow$ single; $\leq 4 \rightarrow$ small; $\leq 8 \rightarrow$ medium; $> 8 \rightarrow$ many_hosts

# comparec: Compare and Recommend

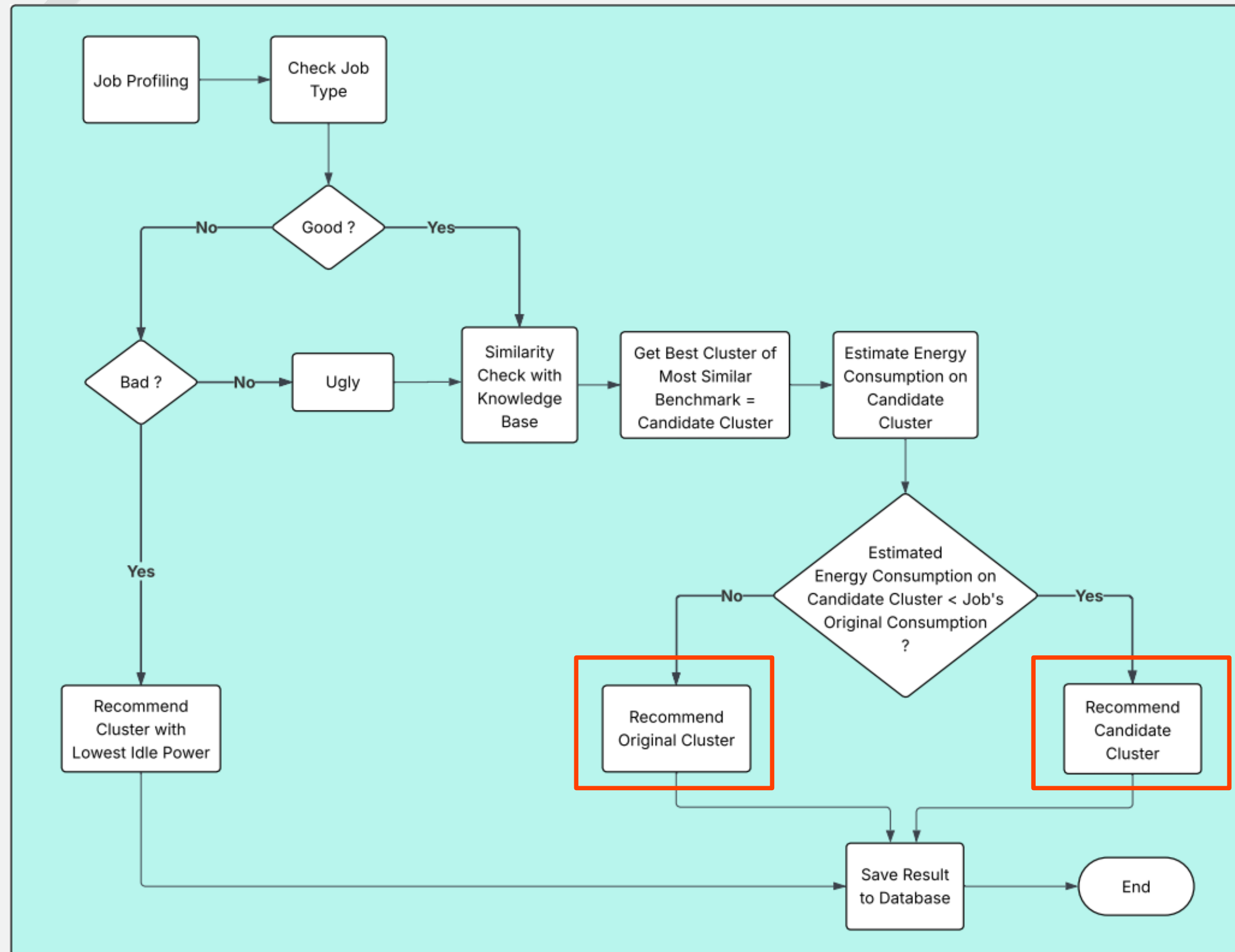


Category	Condition
Bad	Near-idle jobs
Good	High usage jobs with short or medium duration
Ugly	CPU jobs on high performance GPU resources High usage jobs with long duration Low or medium usage jobs Any other unidentified job type

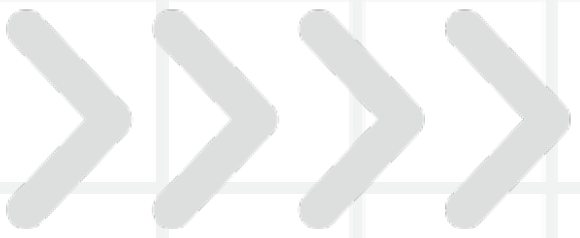
# comparec: Compare and Recommend



# comparec: Compare and Recommend

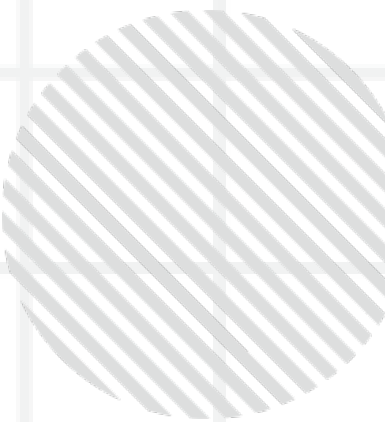


**Recommend** a cluster for a historical (finished) job for its **re-execution**



# QUESTION

How can we **evaluate** the **accuracy** of the recommendations made by **comparec** - the core module ?



# Evaluation of Recommendations



Type	Workloads
PolyBench	2mm (matrix=5000), 3mm (matrix=5500), gemm (matrix=5000), covariance (4000, 8500), durbin (N=900000), fdttd-2d (tmax=200000, nx=2000, ny=2600), jacobild (tsteps=1000000000, n=5000), seidel-2d (nsteps=2500, n=4000)
CPU	sysbench_cpu_prime (prime=5000, multi-thread), sysbench_threads (yields=5M, locks=512), stressng_matrix (20M ops), idle.sh (MPI sleep)
Memory	sysbench_memory (block=1M, total=1200G), stressng_vm (90% memory, 2M ops)
Storage	fiio_seqread (1M block, 10G file), fio_seqwrite (1M block, 10G file), fio_randwrite (4k block, iodepth=128), ioping (20 runs)
MPI	osu_gather (400 runs), osu_reduce (400 runs)

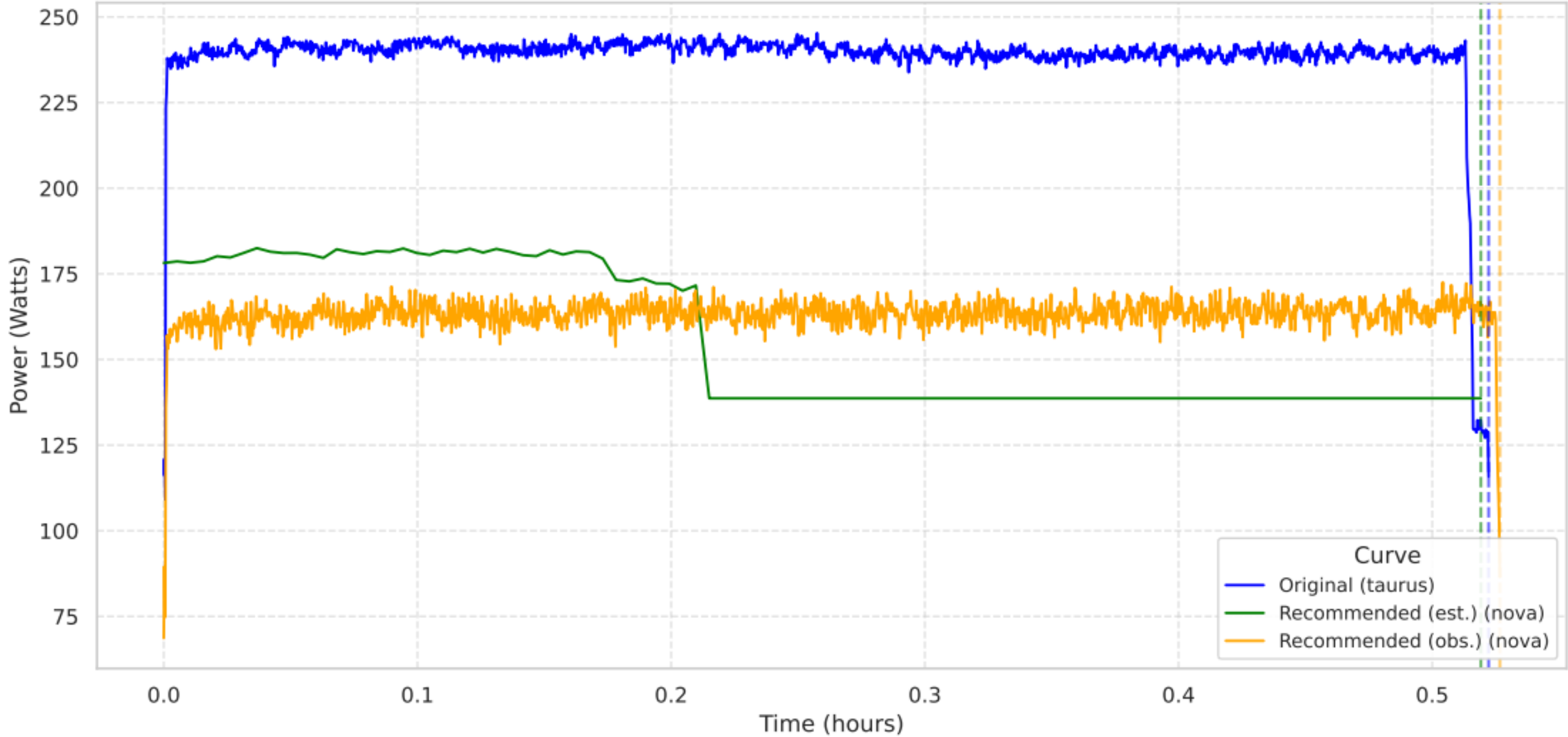
1. Executed a workload on a cluster **chosen by OAR**;
2. Fed workload (1) logs to recommendation system to get a **recommended** cluster for it;
3. Executed workload (1) on the recommended cluster (2), as well as **all other clusters** of the resource provider.



# Evaluation of Recommendations

Recommendation Correct (nova)  
Est. savings: 0.043 kWh | 0.2 min  
Obs. savings: 0.038 kWh | -0.3 min  
Best-Case savings: 0.038 kWh | -0.3 min

2mm\_all\_5000.sh | Job 1927825  
Original: taurus | Recommended: nova | Best: nova



Results on one of the evaluation workloads. Recommended cluster = **best observed** cluster

# Evaluation Results Summary

	Energy		Time	
	kWh	Savings %	h	Savings %
Initial Execution	2.40	0.0	10.78	0.0
Recommended (est.)	1.47	38.8	10.47	2.9
<b>Recommended (obs.)</b>	1.36	43.3	10.29	4.5
Best-Case Execution	1.16	51.7	10.15	5.8

# Evaluation Results Summary

	Energy		Time	
	kWh	Savings %	h	Savings %
Initial Execution	2.40	0.0	10.78	0.0
Recommended (est.)	1.47	38.8	10.47	2.9
<b>Recommended (obs.)</b>	1.36	43.3	10.29	4.5
Best-Case Execution	1.16	51.7	10.15	5.8

Average **relative position score** of recommended clusters : **0.97**  
(0 – worst, 1 – best cluster)

# Evaluation Results Summary

	Energy		Time	
	kWh	Savings %	h	Savings %
Initial Execution	2.40	0.0	10.78	0.0
Recommended (est.)	1.47	38.8	10.47	2.9
<b>Recommended (obs.)</b>	1.36	43.3	10.29	4.5
Best-Case Execution	1.16	51.7	10.15	5.8

Average **relative position score** of recommended clusters : **0.97**  
(0 – worst, 1 – best cluster)

**Initial Execution** – Original clusters ;

**Recommended (est.)** – Recommendation system estimations ;

**Recommended (obs.)** – Observed results on recommended clusters

**Best-Case Execution** – Consumption on most efficient cluster identified during evaluation

# Evaluation Results Summary

	Energy		Time	
	kWh	Savings %	h	Savings %
Initial Execution	2.40	0.0	10.78	0.0
Recommended (est.)	1.47	38.8	10.47	2.9
<b>Recommended (obs.)</b>	1.36	43.3	10.29	4.5
Best-Case Execution	1.16	51.7	10.15	5.8

Average **relative position score** of recommended clusters : **0.97**  
(0 – worst, 1 – best cluster)

**Initial Execution** – Original clusters ;

**Recommended (est.)** – Recommendation system estimations ;

**Recommended (obs.)** – Observed results on recommended clusters ;

**Best-Case Execution** – Consumption on most efficient cluster identified during evaluation

**Observed savings** when recommendations were followed : **83.8 % of the best-case** scenario savings



## CONCLUSION AND FUTURE WORK

1. We presented design of cluster recommendation system with **core module**
2. Evaluations of **comparec** (core module) on Grid'5000 showed an average relative position score of **0.97** (0 – worst, 1 – best cluster).
3. Future work will focus on **enriching knowledge base** for more accurate recommendations and testing with more evaluation workloads.
4. We will **implement** the cluster recommendation system on a **real platform** with real users to test for a few months.



Institut de Recherche en Informatique de Toulouse  
CNRS - Toulouse INP - UT - UT Capitole - UT2



# THANK YOU !

»»»» **QUESTIONS ?**