TRUE NORTH PROGNOSTICS

Optimized Power Management for Air-Cooled Data Centers

A telemetry-driven framework for improving energy productivity

A supplementary framework for evaluating total energy consumed per completed workload in air-cooled environments.

OPM™ overview version

May 2026

What it is: a closed-loop energy productivity layer

- OPM™ is a telemetry-driven control system designed to minimize total energy consumed per unit of useful work.
- It is built for air-cooled data centers where compute, storage and cooling behavior interact.
- The framework applies to compute-intensive and I/O-intensive workloads, including AI training, HPC, simulations, rendering, databases, analytics and transactional systems.
- It supplements PUE and existing thermal guidance rather than replacing them.

OPM™ optimizes computational energy productivity, not facility overhead alone.

Lower

energy per workload

Faster

workload completion

Higher

asset utilization

What it does (and does not do)

What it does

Optimize total energy per workload
Reduce thermal throttling
Reduce fan-power spikes
Reduce vibration-related storage penalties
Improve job completion time

What it does not do

Change power distribution infrastructure
Replace cooling infrastructure
Assume higher setpoints are always better
Aim to improve PUE directly

PUE measures facility overhead. OPM™ optimizes computational energy productivity.

Why facility-only metrics can miss workload energy

What PUE captures

- PUE is useful for measuring facility overhead.
- It compares total facility energy to IT energy.
- It does not measure useful work completed.

What can be hidden

- A lower PUE can coincide with higher energy per workload if runtime rises.
- Thermal stress can trigger leakage, throttling, fan power and storage latency.
- Energy and throughput must be reported together to judge productivity.

The operating target is minimum energy per completed workload - not minimum PUE alone.

Key mechanisms that increase energy per workload

Chip leakage

Leakage power rises with temperature, increasing consumed energy.

DVFS / throttling

Throughput can drop near thermal limits; jobs run longer.

Fan power

Fan power changes nonlinearly with RPM, creating spikes.

Storage effects

HDD latency and I/O retry penalties can rise under vibration.

Telemetry sources used by OPM™

Electrical

On-board voltage and current sensors

Thermal / control

CPU/GPU thermal sensors and DVFS state data

Mechanical / I/O

Fan RPM, vibration signatures and storage I/O counters

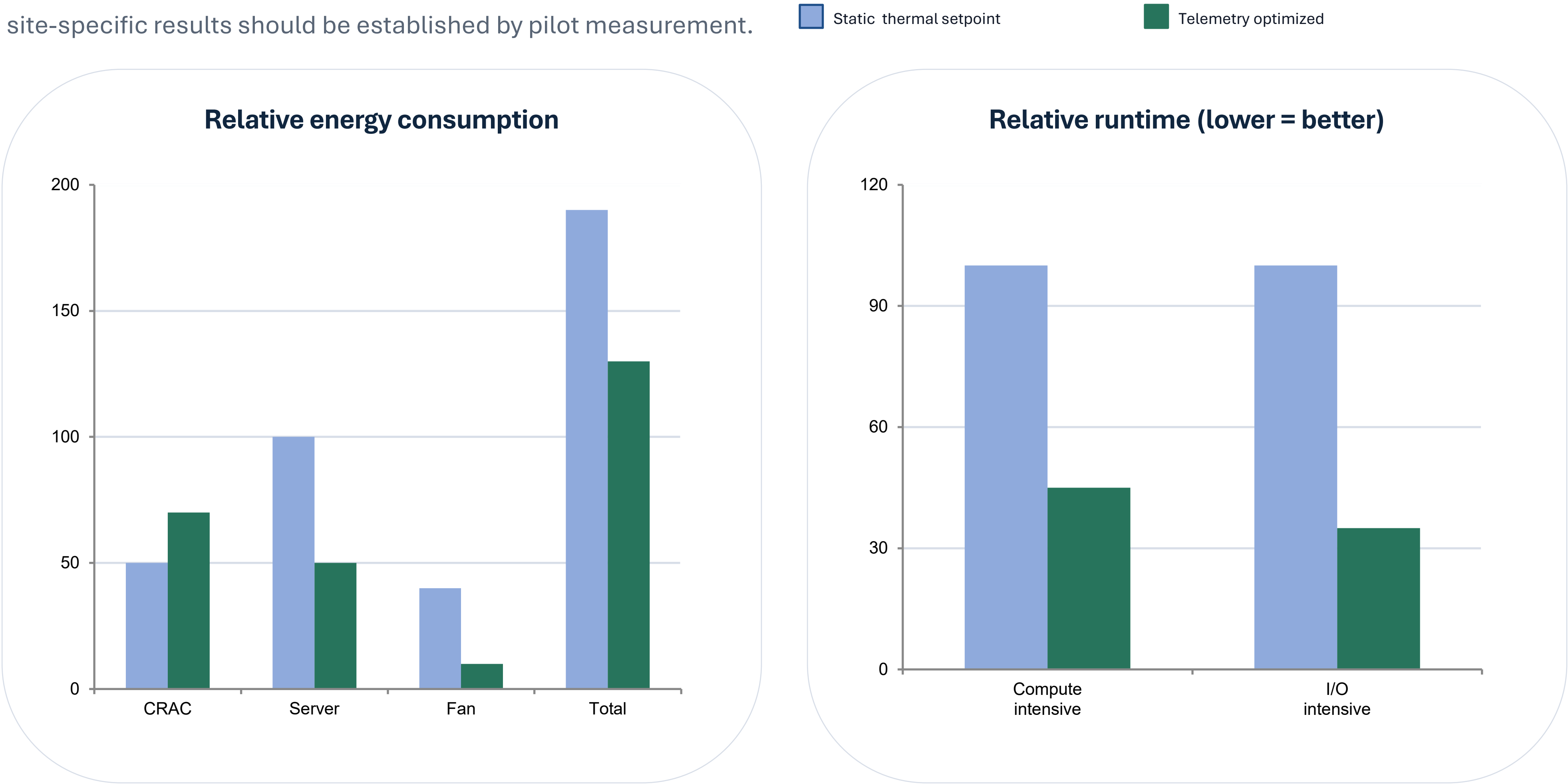
These effects can increase total energy per job even when instantaneous IT power appears lower.

Conceptual comparison of operating modes

Illustrative comparison; site-specific results should be established by pilot measurement.

Compare

Runtime
Fan power
Throttling
Storage latency
Total energy per workload



ASHRAE™ - American Society of Heating, Refrigerating and Air-Conditioning Engineers

Where it delivers the most impact

Compute-intensive

AI training, HPC, simulations and rendering workloads where runtime and throttling matter.

I/O-intensive

Databases, analytics and transactional systems where storage latency affects job energy.

Air-cooled sites

Facilities using CRAC/CRAH systems, including HDD-heavy environments.

OPM™ is most useful where IT behavior, cooling setpoints and workload scheduling interact.

Expected outcomes under suitable conditions

When properly implemented in air-cooled environments with mixed compute and storage loads, OPM™ can deliver:

Reduced

total electricity consumption

Faster

workload completion

Higher

asset utilization

Lower

emissions per unit of work

Savings come from avoiding throttling, reducing fan spikes, limiting leakage escalation, shortening runtime and improving setpoint selection.

Operating benefits

- Deferred infrastructure expansion
- Lower carbon emissions per unit of useful work
- Better reporting of energy and throughput together

Recommended pilot validation framework

1. **Baseline** — Collect telemetry and energy data before intervention.
2. **Controlled intervention** — Change setpoints or controls within agreed boundaries.
3. **Normalize** — Match or normalize workloads before comparing results.
4. **Report** — Publish energy and throughput together.
5. **Accept** — Include reliability and service-level acceptance criteria.

- Pilot measurement should establish site-specific performance before operational adoption.
- The goal is not a generic savings claim; the goal is a measured energy-productivity result.

Selected technical basis

1	**Leakage-aware control**	Workload/cooling management accounts for temperature-driven leakage power.
2	**Fan-speed-aware scheduling**	Scheduling and setpoints consider nonlinear fan energy.
3	**DVFS interaction**	Control choices account for performance and energy tradeoffs.
4	**Storage degradation**	Thermo-mechanical effects can influence HDD latency and I/O rates.
5	**Cold aisle optimization**	Setpoints are optimized against workload energy, not PUE alone.

The common thread: control decisions must include IT energy, workload runtime and cooling behavior together.

Core principle

Cooling setpoints should be optimized for minimum energy per workload, not minimum PUE alone.

- PUE remains useful for facility overhead.
- OPM™ adds a workload-centered lens: energy consumed per useful work completed.
- The pilot objective is to prove that the combined metric improves operating decisions.

Measure energy and throughput together before declaring a setpoint optimal.

Bibliography: energy productivity mechanisms

External scientific publications documenting the physics behind the three parasitic energy-wastage mechanisms: CPU/GPU leakage power, fan motor power and system ambient vibrations.

- "Simware: A Holistic Warehouse-Scale Computer Simulator," Georgia Tech Journal Paper, IEEE Computer Journal (0018-9162/12), 2012. Shows that as data center temperature rises, fan motor power in IT assets rises significantly, wasting energy and distorting facility PUE.
- "Thermo-Mechanical Coupling Induced Performance Degradation in Storage Systems," S. Sondur, K. C. Gross, and K. Kant, 20th IEEE/ACM International Symposium on Cluster, Cloud, and Internet Computing, Melbourne, Australia (May 11-14, 2020).
- "Leakage-Aware Workload and Cooling Management for Improving Server Energy Efficiency," M. Zapater, O. Tuncer, J. L. Ayala, J. M. Moya, K. Vaidyanathan, K. C. Gross, and A. K. Coskun, Transactions on Parallel and Distributed Systems Journal (June 29, 2014).
- "Data Center Ambient Temperature Optimization for Large-Scale Energy Savings," K. C. Gross, Invited Paper and Keynote Industrial Presentation, EcoCloud Annual Symposium on Enterprise Computing Energy Optimization, EPFL, Switzerland (June 2015).
- "Fan-Speed-Aware Scheduling of Data-Intensive Workloads for Optimal Performance and Energy Efficiency of Data Center Assets," C. S. Chan, Y. Jin, Y. K. Wu, K. C. Gross, K. Vaidyanathan, and T. S. Rosing, Proc. International Symposium on Low Power Electronics and Design (ISLPED12), Redondo Beach, CA (Jul 2012).
- "Correcting Vibration-induced Performance Degradation in Enterprise Servers," C. S. Chan, B. Pan, K. C. Gross, K. Vaidyanathan, and T. S. Rosing, Proc. ACM 2013 International Green Metrics Conference, Pittsburgh, PA (Jun 2013).
- "Data Center Cold Aisle Set Point Optimization," Rubenstein and Faist, Microsoft Research, Proc. IEEE ITHERM Conference (June 2014).
- "Effects of Data Center Vibration on Compute System Performance," Julius Turner, Proceedings of USENIX Workshop on Sustainable Information Technology, San Jose, CA (2010). Shows 85% degradation in IO rates for IO-intensive workloads in on-prem data centers.
- "Intelligent Power Monitoring and Management for Enterprise Servers," K. Vaidyanathan and K. C. Gross, 3rd Berkeley IEEE Symposium on Energy Efficient Electronic Systems (E3S-2013), Oct 28-29, 2013.
- "Impact of Dynamic Voltage and Frequency Scaling on n-Tier Application Performance," Georgia Tech and Fujitsu Laboratories, Proc. of ACM Conf. on Timely Results in Operating Systems (TRIOS 2013), Nov 2013.
- "Leakage-Aware and Temperature-Aware Server Control for Improving Energy Efficiency in Data Centers," M. Zapater, J. L. Ayala, J. M. Moya, K. Vaidyanathan, K. C. Gross and A. K. Coskun, International Symposium on Design Automation and Test in Europe (DATE-2013), Grenoble, France (Apr 2013).
- "Ambient Temperature Optimization for Enterprise Servers: Key to Large-Scale Energy Savings," K. Vaidyanathan, K. C. Gross, S. Sondur, 4th Berkeley IEEE Symposium on Energy Efficient Electronic Systems (S3E-2015), Berkeley, CA (Oct 28-29, 2015).