REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

ESCALONAMENTO CIENTE DE DIAFONIA TÉRMICA E DVFS PREDITIVO BASEADO EM MODELO PARA CLUSTERS DE ACELERADORES DE IA ESPACIAIS EM ALTO VÁCUO: MODELAGEM MULTIFÍSICA APROXIMADA, CONTROLE PREDITIVO E SIMULAÇÃO NUMÉRICA

THERMAL-CROSSTALK-AWARE SCHEDULING AND MODEL PREDICTIVE DYNAMIC VOLTAGE AND FREQUENCY SCALING FOR SPACEBORNE MULTI-CHIP AI CLUSTERS IN DEEP VACUUM: LUMPED MULTIPHYSICS MODELING, PREDICTIVE CONTROL, AND NUMERICAL SIMULATION

PROGRAMACIÓN CONSCIENTE DE DIAFONÍA TÉRMICA Y DVFS PREDICTIVO BASADO EN MODELOS PARA CLÚSTERES DE ACELERADORES DE IA ESPACIALES EN ALTO VACÍO: MODELADO MULTIFÍSICO APROXIMADO, CONTROL PREDICTIVO Y SIMULACIÓN NUMÉRICA

Dafnhê Thor Castro de Martins Barros
Master's Student in the Graduate Program in Control and Automation Engineering (ProPECAut), Federal Institute of Espírito Santo (IFES), Serra, ES, Brazil
ORCID: 0009-0004-9840-0257
E-mail: dafnhe.barros@msn.com

Flávio Barcelos Braz da Silva
D.Sc. in Electrical Engineering, Permanent Professor in the Graduate Program in Control and Automation Engineering (ProPECAut), Federal Institute of Espírito Santo (IFES), Serra, ES, Brazil
ORCID: 0000-0002-4752-9742
E-mail: flavio.silva@ifes.edu.br

Abstract

High-density heterogeneous artificial intelligence (AI) computing clusters onboard small satellites in Low Earth Orbit (LEO) face severe electrothermal constraints. Gas-mediated convective heat transfer is neglected in the adopted high-vacuum model ($h \approx 0$), confining thermal rejection exclusively to internal solid conduction and non-linear Stefan-Boltzmann radiation to deep space ($q \propto T^4$). In multi-chip modules (MCM), heavy neural inference workloads induce severe lateral conductive coupling between adjacent silicon dies (thermal crosstalk). This paper investigates a model-based predictive thermal management strategy, denoted TCAS-DVFS (Thermal-Crosstalk-Aware Scheduling and Model Predictive Dynamic Voltage and Frequency Scaling), formulated from lumped thermodynamic principles for multi-die spaceborne accelerators. The proposed architecture integrates: (i) online state estimation at the simulated sampling rate via an Extended Kalman Filter (EKF) operating over a non-linear 5×5 Compact Thermal Model (CTM) with exact analytical subthreshold leakage and radiative Jacobians; (ii) a discrete linear time-varying Model Predictive Controller (LTV MPC-DVFS) that penalizes violations of a nominal 83 °C soft thermal cap via decoupled spatio-temporal slack variables, solved via a warm-started ADMM splitting iteration paired with an explicit shifted-tail power-projected fallback for non-convergent iterations; (iii) a physical radiative and conductive

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

power-budget closure with transient enthalpy buffering derived from a synthetic analytical orbital radiative forcing model; and (iv) a topological task allocator relaxed over the semidefinite cone S_+^{17} via Semidefinite Programming (SDP). In a three-orbit numerical co-simulation of a 500-km circular LEO scenario (17,004.0 s, 34,009 time samples, 4 processing dies), TCAS-DVFS produced zero observed junction temperature violations above the adopted 85 °C simulation thermal limit (0 s versus 200.5 s for binary Thermal Throttling at 85/75 °C and 14.0 s for Space-PID), attenuated peak inter-die thermal gradients by 78.7% (reducing from 9.42 to 2.01 K, and mean gradients by 83.1% from 5.26 to 0.89 K), and maintained an aggregate throughput of 124.62 FPS with an energy-delay product proxy of 7.45 kJ·s/frame. Functional ablation indicates a dominant contribution from the predictive control formulation in damping gradients and bounding temperatures, whereas topological SDP scheduling matches the best discrete permutation among the 4! = 24 candidates after hyperplane rounding (142/142 epochs) despite partial ADMM numerical convergence, providing limited net thermal benefit in compact symmetric arrangements (N = 4 dies). Descriptive parametric sensitivity was evaluated under a 30-sample stratified Latin Hypercube Sampling (LHS) campaign (±20% perturbation across the four conductive and capacitive network parameters R_ij, R_r, c_i, and C_rad), yielding zero observed violations (0/30 samples > 85 °C), where die-to-radiator thermal resistance (R_r) exhibited the strongest monotonic association with T_max (Spearman ρ = +0.995). Indicative reliability proxy analyses demonstrate that electromigration yields no net life extension (relative Black damage ratios of 0.97 and 0.83), whereas cyclic thermal stress severity is substantially damped. The reported evidence is limited to numerical simulation of the stated lumped-parameter model and does not constitute hardware or flight qualification.

Keywords: Control applications; Model predictive control; Thermal crosstalk; Dynamic voltage and frequency scaling; Spaceborne computing; Orbital vacuum thermodynamics.

Resumo

Clusters heterogêneos de alta densidade voltados à inferência de inteligência artificial (IA) a bordo de pequenos satélites em Órbita Baixa da Terra (LEO) enfrentam severas restrições eletrotérmicas. No vácuo espacial, a convecção gasosa é desprezada no modelo adotado ($h \approx 0$), restringindo a rejeição térmica estritamente à condução sólida interna e à radiação não linear de Stefan-Boltzmann para o espaço profundo ($q \propto T^4$). Em módulos multi-chip (MCM), rajadas intensivas de inferência induzem forte acoplamento condutivo lateral entre matrizes adjacentes de silício (diafonia térmica espacial). Este artigo desenvolve e avalia uma estratégia preditiva de gerenciamento térmico baseada em modelo, denominada TCAS-DVFS (Thermal-Crosstalk-Aware Scheduling and Model Predictive Dynamic Voltage and Frequency Scaling), formulada a partir de primeiros princípios termodinâmicos concentrados. A arquitetura proposta integra: (i) estimação de estados em tempo de simulação via Filtro de Kalman Estendido (EKF) com formulação estabilizada de Joseph sobre uma rede térmica compacta 5×5 (CTM) com Jacobianos analíticos exatos; (ii) um Controlador Preditivo Baseado em Modelo Linear Variante no Tempo (LTV MPC-DVFS) discreto com variáveis de folga espaço-temporais desacopladas penalizando um teto suave nominal de 83 °C, resolvido via divisão de operadores ADMM com inicialização a quente e política determinística explícita de fallback de cauda deslocada com projeção de potência; (iii) fechamento de orçamento físico de potência radiativa e condutiva com buffer entalpico transiente sob modelo analítico de forçamento radiativo orbital; e (iv) um alocador topológico de tarefas relaxado sobre o cone semidefinido S_+^{17} via Programação Semidefinida (SDP). Em co-simulação numérica de 3 órbitas completas em LEO (17.004,0 s, 34.009 passos, 4 matrizes ativas), o TCAS-DVFS resultou em zero violações da temperatura de junção acima do limite de simulação de 85 °C (0 s frente a 200,5 s do Throttling 85/75 °C e 14,0 s do Space-PID), atenuou gradientes térmicos espaciais de pico em 78,7% (de 9,42 K para 2,01 K, e médios em 83,1% de 5,26

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

para 0,89 K) e sustentou vazão de 124,62 FPS com produto energia-atraso de 7,45 kJ·s/quadro. A ablação funcional indica contribuição dominante do controle preditivo, enquanto o escalonador SDP empata com a enumeração exata das 4! = 24 permutações após arredondamento (100% de acerto em 142 épocas), com benefício líquido marginal em arranjos simétricos compactos (N = 4). A sensibilidade paramétrica descritiva sob Amostragem por Hipercubo Latino (LHS estratificada, n=30, variação de ±20% nos 4 parâmetros passivos R_ij, R_r, c_i, C_rad) confirmou zero violações acima de 85 °C (0/30 amostras), dominada pela resistência die-radiador (R_r, Spearman ρ = +0,995). Análises de confiabilidade indicativa revelam que a eletromigração não apresenta extensão de vida útil (razões de dano relativo de Black de 0,97 e 0,83), enquanto a severidade de fadiga térmica cíclica é substancialmente amortecida. As evidências limitam-se à simulação numérica do modelo de parâmetros concentrados, sem constituir qualificação física de voo.

Palavras-chave: Clusters de IA espaciais; Controle preditivo baseado em modelo; Diafonia térmica espacial; DVFS preditivo; Filtro de Kalman Estendido; Termodinâmica no vácuo.

Resumen

Los clústeres heterogéneos de computación de inteligencia artificial (IA) a bordo de pequeños satélites en Órbita Terrestre Baja (LEO) enfrentan severas restricciones electrotérmicas. En el alto vacío espacial, la transferencia convectiva es despreciable (h ≈ 0), limitando la disipación térmica a la conducción sólida interna y a la radiación no lineal de Stefan-Boltzmann hacia el espacio profundo ($q \propto T^4$). En módulos multichip (MCM), las ráfagas intensivas de inferencia neural provocan un fuerte acoplamiento conductivo lateral entre matrices de silicio adyacentes (diafonía térmica espacial). Este artículo desarrolla y evalúa una estrategia predictiva basada en modelos, denominada TCAS-DVFS (Thermal-Crosstalk-Aware Scheduling and Model Predictive Dynamic Voltage and Frequency Scaling), formulada a partir de principios termodinámicos concentrados. La arquitectura propuesta integra: (i) estimación de estado en tiempo de simulación mediante un Filtro de Kalman Extendido (EKF) con formulación estabilizada de Joseph sobre un modelo térmico compacto 5×5 (CTM); (ii) un Controlador Predictivo Lineal Variante en el Tiempo (LTV MPC-DVFS) discreto resuelto mediante división tipo ADMM con inicio en caliente y política determinista explícita de fallback; (iii) balance físico de potencia radiativa y conductiva con amortiguamiento entálpico transitorio; y (iv) un asignador topológico de tareas relajado en el cono semidefinido S_+^{17} mediante Programación Semidefinida (SDP). En cosimulación numérica de 3 órbitas LEO (17.004,0 s, 34.009 muestras, 4 matrices), TCAS-DVFS logró cero violaciones térmicas sobre el límite de simulación de 85 °C (0 s frente a 200,5 s de Throttling y 14,0 s de Space-PID), atenuó los gradientes térmicos espaciales de pico en un 78,7% (de 9,42 K a 2,01 K, y medios en 83,1% de 5,26 a 0,89 K) y mantuvo un rendimiento de 124,62 FPS con un producto energía-retraso de 7,45 kJ·s/cuadro. Bajo muestreo estratificado de hipercubo latino (LHS, n=30, ±20% en los 4 parámetros pasivos), se obtuvo cero violaciones sobre 85 °C, dominado por la resistencia die-radiador (R_r, Spearman ρ = +0,995). Los análisis de confiabilidad indican que la electromigración no ofrece extensión neta de vida útil (daño relativo de Black de 0,97 y 0,83), mientras que la fatiga térmica cíclica se atenúa sustancialmente. Las evidencias se restringen a simulación numérica del modelo concentrado.

Palabras clave: Clústeres espaciales de IA; Control predictivo basado en modelos; Diafonía térmica espacial; DVFS predictivo; Filtro de Kalman Extendido; Termodinámica del vacío.

1. Introduction

1.1 Orbital Edge Computing and Spaceborne AI

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

Modern Earth Observation (EO) satellite missions, including wildfire monitoring, maritime surveillance, disaster response, and Synthetic Aperture Radar (SAR) imaging, produce massive data streams operating at multi-gigabit-per-second rates in Low Earth Orbit (LEO) (Denby and Lucia 939; Garrett, Roffe, and George 5481; H. Chen et al. 103969). Traditionally, remote sensing platforms have relied on the conventional store-and-forward architecture, wherein raw payload data are captured, buffered in solid-state recorders, and downlinked during brief 8-to-12-minute ground-station passes per orbit. This approach incurs severe latency penalties, frequently exceeding 4 to 12 hours.

Edge artificial intelligence (AI) onboard small satellites mitigates this downlink bottleneck by processing, compressing, and triaging sensor data directly in orbit using commercial off-the-shelf (COTS) neural accelerators, multicore ARM CPUs, and embedded GPUs (Garrett, Roffe, and George 5481; Sherwood et al. 1-390). In particular, automated on-orbit cloud segmentation enables the autonomous discarding of up to 67% of optically obstructed imagery prior to transmission (King et al. 3826), thereby conserving scarce communication bandwidth for high-priority mission products (Giuffrida et al. 1).

1.2 Deep Vacuum Thermodynamics and Thermal Crosstalk

Unlike terrestrial data centers that leverage forced air convection (h ≈ 20–100 W/(m²K)) or pumped liquid cooling loops (h ≈ 1,000–10,000 W/(m²K)) (Bergman et al. 45), small spacecraft operating in the high vacuum of LEO ($P < 10^{-4}$ Pa, free molecular flow with Knudsen number Kn >> 10) (Gilmore 15) exhibit negligible convective gas exchange. Consequently, gas-mediated convective heat transfer is neglected in the adopted high-vacuum model:

$$h \equiv 0 \tag{1}$$

Thermal rejection from energized silicon junctions is restricted to two coupled mechanisms (ECSS 12; Gilmore 45):

1. Internal solid conduction: heat dissipates from active junctions across thermal interface materials (TIM) into a shared cold plate or metallic carrier (e.g., AlN or Al 6061-T6, with k ≈ 170–200 W/(m·K)). Direct radiative exchange between facing

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

die side-walls is negligible (< 10^{-3} W due to microscopic lateral areas $A_{lat} \approx 10^{-5}$ m² and small temperature differences).

2. External radiation to deep space: structural heat reaches outer passive radiators and is radiated to the cold cosmic background ($T_{space} \approx 2.73$ K) governed by the non-linear Stefan-Boltzmann law ($q_{rad} \propto \varepsilon \sigma T^4$) (Pederson et al. 1).

In compact small-satellite form factors (e.g., 6U–12U CubeSats), volume and mass constraints dictate packaging multiple System-on-Module (SoM) accelerators on a single carrier. Simultaneous deep neural inference workloads dissipate between 10 and 25 W per die (Reuther et al. 1; Kang and Park 3840), demanding 60 to 100 W from the electrical power system (EPS). Crucially, lateral conductive flux propagating through the carrier raises the base temperature of neighboring idle or lightly loaded dies by 10 to 18 K within tens of seconds (S. Chen et al. 1; Wang et al. 1). This mutual conductive coupling, termed spatial thermal crosstalk, induces mutual thermal interference and severely degrades uncoordinated, decentralized governors.

1.3 Deficiencies of Terrestrial Thermal Governors

Conventional reactive algorithms, such as the Linux ondemand governor (Pallipadi and Starikovskiy 1) and classical Single-Input Single-Output (SISO) PID loops, fail in orbital clusters due to three structural factors (Brooks and Martonosi 171; Skadron et al. 94; Hoffmann and Fröhlich 493):

1. Large thermal inertia: passive radiators ($C_{rad} \approx 1{,}300$ J/K) introduce dominant time constants of tens of minutes (1,500 to 2,100 s), causing purely reactive policies to act with critical delay.

2. Limit-cycle saw-tooth oscillations: binary frequency throttling upon crossing emergency alarm thresholds triggers repetitive temperature fluctuations ($\Delta T \approx 10$ K).

3. Spatial gradient amplification: uncoordinated DVFS across asymmetric workloads exacerbates instantaneous temperature differentials across adjacent dies, accelerating low-cycle thermomechanical shear fatigue in Ball Grid Array (BGA) solder interconnects (Manson 1).

1.4 Research Question and Contributions

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

The pioneering work of S. Chen et al. (1) established thermal-crosstalk-aware task scheduling in spaceborne AI clusters. Building on this domain, our work investigates the following research question:

Research Question: "Can a model-based coordinated MPC-DVFS policy reduce thermal-threshold exceedance and inter-die temperature gradients in a simulated four-die orbital AI cluster while maintaining competitive inference throughput relative to reactive thermal governors under identical modeled workload and environmental conditions?"

To address this question through model-based numerical simulation, this paper presents the following contributions within the scope of the modeled multi-physics system:

1. Non-linear State-Space Electrothermal Modeling: we formulate a 5×5 compact thermal network (CTM) capturing conductive coupling matrices, linearized Stefan-Boltzmann radiator dynamics, and non-linear subthreshold leakage positive feedback, integrated with an Extended Kalman Filter (EKF) observer under stochastic sensor noise.

2. Constrained Predictive Control (LTV MPC-DVFS): we formulate a discrete linear time-varying Model Predictive Controller that solves a condensed quadratic programming problem via a warm-started ADMM-style splitting iteration, coordinating operating frequencies to track inference workloads, enforce a nominal 83 °C limit, and suppress spatial gradients.

3. Orbital Radiative Power Budget Surrogate: we derive a predictive moving-average budget for solar flux, terrestrial albedo, and Earth infrared radiation to establish a temperature-referenced power envelope under eclipse transitions.

4. Numerical Sensitivity and Methodological Demarcation: we execute a three-orbit numerical co-simulation, functional ablation, a descriptive 30-sample stratified Latin Hypercube Sampling (LHS) sensitivity analysis across the four passive conductive and capacitive network parameters (R_{ij}, R_r, c_i, C_{rad}), and physics-of-failure reliability proxy evaluations.

Table 1 – Methodological Demarcation of Study Elements.

Methodological Category	Specific Elements in this Study

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

Input Physical Parameters	Conductive resistances Rij, die-to-radiator resistance Rr, die capacitances ci, radiator capacitance Crad, optical emissivity ε, solar absorptivity αs (Table 2).
Model Hypotheses & Structure	Lumped-parameter 5×5 CTM; gas conduction neglected (h ≈ 0); linearized radiative state coupling; planar symmetry.
Synthetic Workload Inputs	Continuous activity factors αi(t) and nominal requested frequencies fnom,i(t) across 4 mission workloads.
Model Environmental Forcing	Direct solar flux Qsol(t), Earth albedo Qalb(t), and terrestrial infrared QIR(t) across a 500-km circular orbit.
Calculated State Variables	Physical silicon junction temperatures Ti(t) (i=1..4) and radiator core temperature Trad(t).
Estimated State Variables	EKF state estimates $\hat{T}i(t)$ and $\hat{T}rad(t)$ reconstructed from noisy sensor measurements (σs = 0.25 K).
Actuation Control Variables	Operating clock frequencies f(t) = [f1..f4]T ∈ [0.30, 1.30] GHz and switched voltages Vi(t).
Direct Numerical Outputs	Maximum temperature Tmax, spatial gradient ΔTesp(t), violation time above 85 °C (t>85), and total energy E.
Evaluated Computational Proxies	Aggregate throughput proxy FPS, energy-delay product proxy Et_frame, relative Black electromigration damage ratio, and ASTM rainflow cycle distribution.
Unmeasured Physical Quantities	Physical component service life, BGA viscoplastic solder strain, internal package micro-gradients, space radiation effects (TID/SEE), and flight qualification.

Source: Prepared by the authors (2026).

Table 1 establishes the transparent methodological boundaries of this investigation, classifying the physical lumped hypotheses, numerical inputs, calculated states, evaluated computational proxies, and physically unmeasured laboratory properties (such as TVAC chamber testing or hardware-in-the-loop qualification).

2. Orbital Environmental Physics and Heat Balance

We consider a synthetic circular low Earth orbit thermal scenario at altitude h = 500 km (R = R_E + h = 6,871 km), with analytical orbital period T_orb = 5,668.14 s (≈ 94.47 min). Defining the orbital phase as φ = mod(t, T_orb)/T_orb, the illuminated orbital fraction is φ_sol = 1 - arcsin(R_E/R)/π under the conservative assumption of solar beta angle β = 0° (which maximizes eclipse duration and thermal shock amplitudes).

2.1 External Orbital Radiative Fluxes

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

With orbital solar noon at φ_m = φ_sol/2 and a smooth illumination operator I_sun(t) = [1 + exp(120(φ - φ_sol))]⁻¹, the absorbed radiative fluxes on a radiator of area A_rad = 0.12 m² (Figure 1) are given by:

$$\cos\theta_{ef}(t) = max[0, \cos(\frac{\pi(\varphi-\varphi_m)}{2\varphi_m})\cos 30^\circ \cos 15^\circ] \quad (2)$$

$$Q_{sol}(t) = \mathbb{I}_{sun}(t) S_{solar} \alpha_s A_{rad} \cos\theta_{ef}(t) \quad (3)$$

$$Q_{alb}(t) = \mathbb{I}_{sun}(t) S_{solar} a_E \alpha_s A_{rad} F \cos\theta_{ef}(t) \quad (4)$$

$$Q_{IR}(t) = \varepsilon\sigma A_{rad} F T_E^4(t) \quad (5)$$

$$T_E(t) = 255 - 15\sin^2(\pi\frac{\varphi-\varphi_m}{\varphi_{sol}}) \quad (6)$$

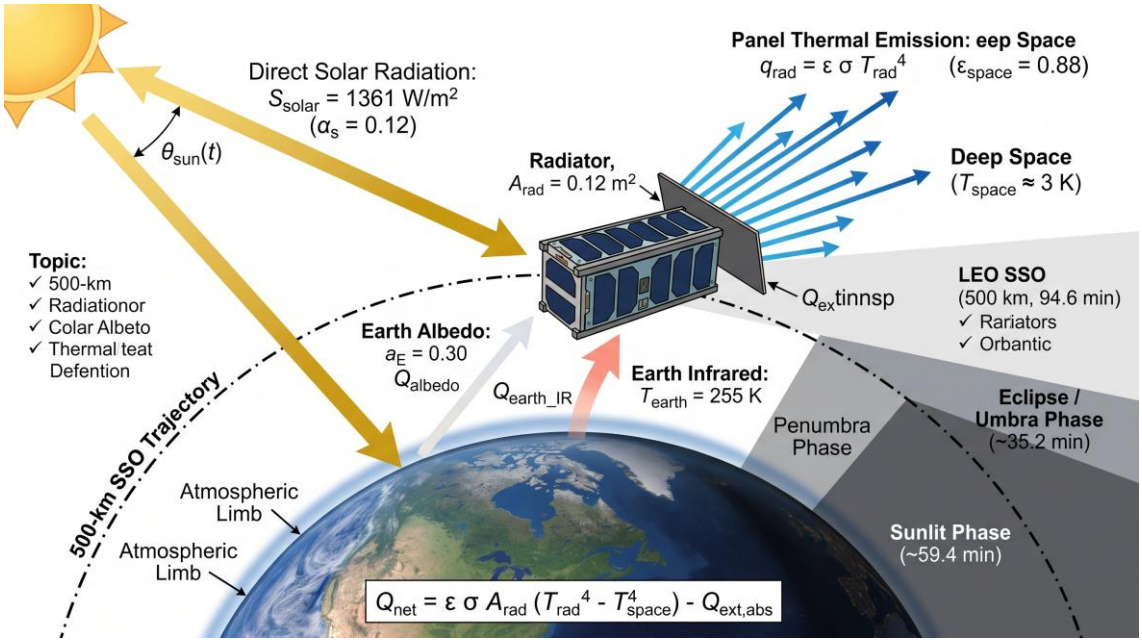
where S_solar = 1,361 W/m² is the solar constant, a_E = 0.30 is the mean Earth albedo, F = 0.42 is the Earth-to-plate radiation view factor, α_s = 0.12 is solar absorptivity, ε = 0.88 is infrared emissivity, and σ = 5.67 × 10⁻⁸ W/(m²K⁴) is the Stefan-Boltzmann constant. Earth albedo is modeled proportionally to the illuminated visible fraction via cos(θ_ef), a standard conservative approximation in spacecraft thermal engineering. The total external flux Q_ext(t) = Q_sol(t) + Q_alb(t) + Q_IR(t) varies between 8.34 W (in eclipse) and 29.09 W (solar noon), with an orbital mean of 16.84 W.

Figure 1 – Synthetic 500 km Circular LEO Orbital Geometry: Direct Solar Flux, Earth Albedo, Infrared Emission, Eclipse Cone, and Deep Space Radiation Rejection.

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34



Source: Prepared by the authors (2026).

Figure 1 illustrates the synthetic circular orbit scenario at 500 km altitude in LEO. The diagram shows direct solar irradiance, diffuse terrestrial albedo, planet infrared flux, the orbital shadow geometry during eclipse transit, and deep-space passive radiative heat rejection from the spacecraft outer radiator panel toward the 3 K cosmic background.

2.2 Passive Radiator Lumped Dynamics

The thermal energy conservation equation for the radiator panel of mass m_rad = 1.45 kg and specific heat c_p,rad = 896 J/(kg·K) (C_rad = m_rad c_p,rad = 1,299.2 J/K) is governed by:

$$C_{rad} \frac{dT_{rad}}{dt} = \sum_{i=1}^{N} \frac{T_i - T_{rad}}{R_r} + Q_{ext}(t) - \varepsilon \sigma A_{rad} (T_{rad}^4 - T_{space}^4) \quad (7)$$

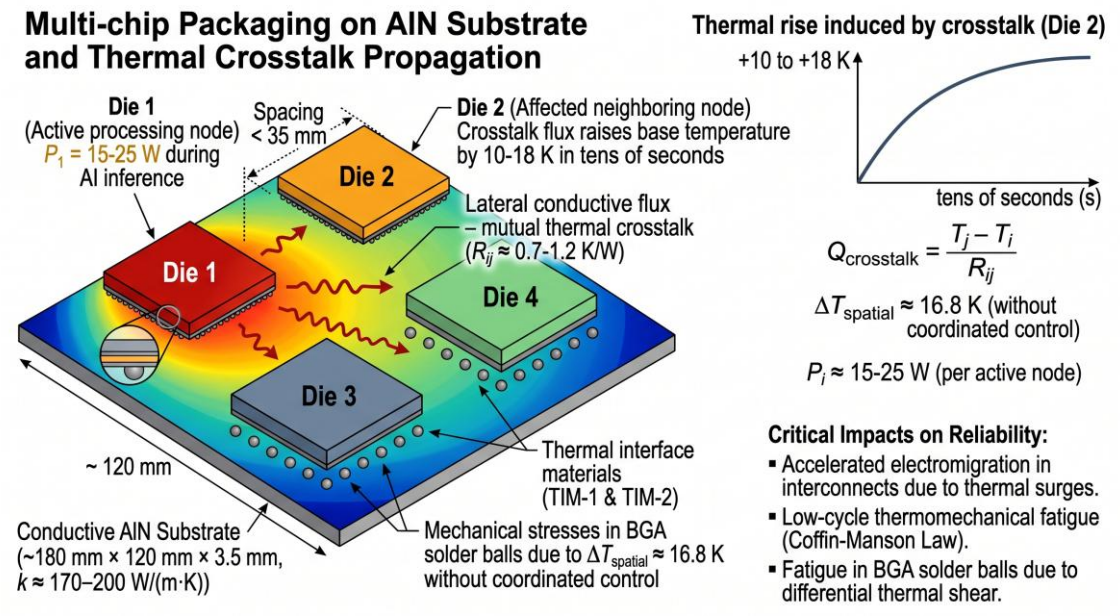
where R_r = 2.10 K/W represents the equivalent solid conductive thermal resistance between each die node and the radiator core.

Figure 2 – Internal Solid Conductive Coupling and Lateral Thermal Crosstalk in a Multi-Chip Carrier.

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34



Source: Prepared by the authors (2026).

Figure 2 illustrates the physical cross-section and 3D architectural layout of the multi-chip module (MCM) in deep vacuum. Active silicon junctions dissipate heat vertically through thermal interface materials (TIM) into the cold plate and laterally through inter-die conductive resistances Rij, raising neighboring die base temperatures and inducing thermomechanical shear stress.

3. State-Space Electrothermal Modeling

3.1 Silicon Power Dissipation and Leakage Feedback

Each silicon die (Figure 2) dissipates electrical power partitioned into dynamic switching and subthreshold leakage components:

$$V_i = \text{clip}(V_{min} + \beta_v(f_i - f_{min}), V_{min}, V_{max}) \quad (8)$$

$$P_{dyn,i} = \alpha_i C_{eff} V_i^2 f_i \quad (9)$$

$$P_{leak,i} = P_{ref}(\frac{V_i}{V_0})^{1.5}(\frac{T_i}{T_0})^2 \exp[\frac{E_a}{k_B}(\frac{1}{T_0} - \frac{1}{T_i})] \quad (10)$$

where total power is P_i = P_dyn,i + P_leak,i, with α_i denoting active workload factor, C_eff = 24.5 nF the effective switching capacitance, and β_v = (V_max - V_min)/(f_max - f_min). Leakage power P_leak,i exhibits a quadratic-exponential

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

dependence on absolute junction temperature T_i, introducing non-linear positive electrothermal feedback (∂P_leak/∂T_i > 0) and a risk of thermal runaway in the absence of active intervention. Nominal parameters reflect commercial 16/12 nm FinFET embedded AI accelerators (Table 2). The thermal threshold T_cap = 85.0 °C is adopted as a typical operational qualification limit in small-satellite payload engineering guidelines (ECSS-E-ST-31C) to maintain safety margins for commercial off-the-shelf electronic components in space missions, whereas 125.0 °C represents the physical upper boundary of validity for standard FinFET subthreshold leakage models documented in solid-state literature (Kang and Park 3840; Skadron et al. 94). Operations above 125 °C are strictly tagged as unphysical mathematical extrapolations (†) and excluded from quantitative comparisons.

3.2 Compact Thermal Network (5×5 CTM)

The planar thermal coupling across an array of N = 4 dies mounted on a shared AlN substrate is modeled by a lumped-parameter compact thermal network:

$$c_i \frac{dT_i}{dt} = P_i(t) - \sum_{j \neq i} \frac{T_i - T_j}{R_{ij}} - \frac{T_i - T_{rad}}{R_r} \quad (11)$$

where c_i = 18.5 J/K is the die thermal capacitance. In compact matrix notation:

$$C\dot{T} = P - (G_{crosstalk} + G_{rad})T + G_{rad}T_{rad}\mathbf{1}_N \quad (12)$$

where C = diag(c_1, ..., c_N), G_rad = R_r^{-1} I_N, and G_crosstalk ∈ $\mathbb{R}$^{N×N} is the symmetric conductive Laplacian matrix:

$$[G_{crosstalk}]_{ij} = \begin{cases} -1/R_{ij}, & i \neq j \\ \sum_{k \neq i} 1/R_{ik}, & i = j \end{cases} \quad (13)$$

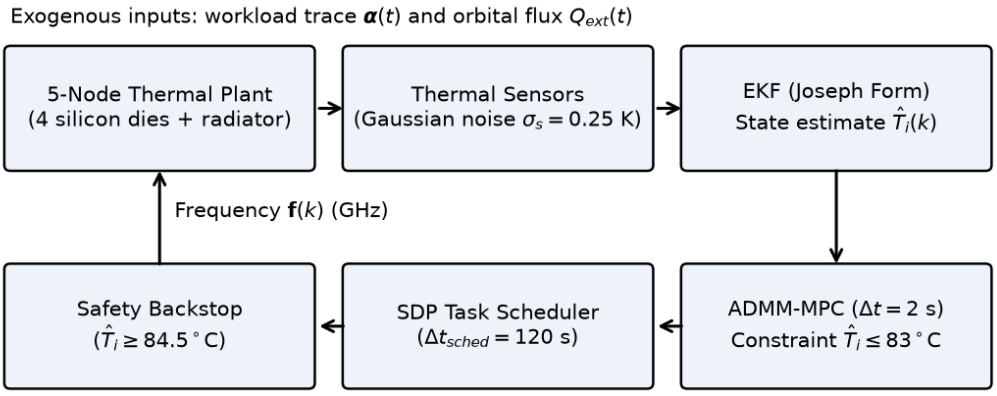
Modal decomposition of the die subsystem yields fast thermal time constants of 5.8, 7.5, 7.7, and 38.8 s, contrasting with the slow radiator time constant (1,500 to 2,100 s). This pronounced two-time-scale separation provides thermodynamic justification for treating radiator temperature as a quasi-stationary state within the short-horizon predictive controller.

Figure 3 – Closed-Loop Control Architecture of the Integrated TCAS-DVFS Framework.

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34



Source: Prepared by the authors (2026).

Figure 3 presents the block diagram of the closed-loop TCAS-DVFS control framework. The system operates on a multi-rate scheme: continuous plant integration and EKF state estimation at 0.5 s, discrete LTV MPC-DVFS quadratic optimization via ADMM at 2.0 s, dynamic radiative power budgeting from orbital ephemerides, and topological SDP task allocation at 120 s.

4. Integrated TCAS-DVFS Control Architecture

The closed-loop control architecture (Figure 3) operates on a multi-rate scheme: continuous physical plant integration and EKF state filtering operate at Δt_plant = 0.5 s, while model predictive control updates execute every k_sub = 4 steps (Δt_MPC = 2.0 s), and topological task allocation updates occur at Δt_sched = 120 s.

4.1 State Estimation via Extended Kalman Filter

The extended state vector $x = [T_1, T_2, T_3, T_4, T_rad]^T \in \mathbb{R}^5$ is measured under zero-mean additive white Gaussian sensor noise: $y = x + v$, $v \sim \mathcal{N}(0, \sigma_s^2 I_5)$, with standard deviation $\sigma_s = 0.25$ K (Zhang and Srivastava 472). State propagation and covariance updates employ the stabilized Joseph form (Simon 1):

$$\hat{x}^- = \hat{x}_{k-1} + \Delta t_{\text{plant}} f(\hat{x}_{k-1}, u_{k-1}) \tag{14}$$

$$\boldsymbol{\Phi} = I_5 + \Delta t_{\text{plant}} F + \frac{1}{2}(\Delta t_{\text{plant}} F)^2 \tag{15}$$

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

$$P^- = \boldsymbol{\Phi} P \boldsymbol{\Phi}^T + Q\Delta t_{\text{plant}} \tag{16}$$

$$K = P^-(P^- + R)^{-1}, \hat{x} = \hat{x}^- + K(y - \hat{x}^-) \tag{17}$$

$$P = (I_5 - K)P^-(I_5 - K)^T + KRK^T \tag{18}$$

where process noise covariance is Q = diag(0.02 I_4, 0.05), measurement noise covariance is R = σ_s² I_5, initial error covariance is P_0 = diag(I_4, 2), and Jacobian F = ∂f/∂x linearizes the Stefan-Boltzmann radiator flux (g_rad,lin = 4 ε σ A_rad T̂_rad³). In closed-loop co-simulation, the observer achieves an average junction temperature tracking RMSE below 0.12 K.

4.2 Predictive Radiative Power Budget Surrogate

Using onboard orbital ephemerides over the prediction horizon (H_p = 30 steps, 60.0 s), we compute a temperature-referenced radiative power-budget surrogate P_bud(k) used as an MPC constraint surrogate to prevent the radiator from driving past steady-state equilibrium:

$$P_{bud}(k) = max\{15, \varepsilon\sigma A_{rad}[(T_{seg} + 273.15)^4 - T_{space}^4] - \overline{Q}_{ext}(k)\} \tag{19}$$

where T_seg = 83 °C and Q̄_ext(k) is the predicted moving-average environmental flux over the lookahead horizon.

4.3 Topological Task Scheduler (TCAS)

At task allocation epochs (Δt_sched = 120 s), M = 4 incoming neural inference workloads are assigned to the N = 4 dies to minimize conductive crosstalk weighted by the steady-state impedance matrix Φ = (G_crosstalk + G_rad)⁻¹:

$$min_X \sum_{i=1}^{N} \quad \sum_{j=1}^{N} \quad \Phi_{ij} P_i(X) P_j(X) + \mu \sum_{i=1}^{N} \quad (T_i - \overline{T})^2 \tag{20}$$

subject to bijective task assignment constraints. Vectorizing the binary assignment decision matrix into x = vec(X) ∈ {0, 1}^{N²} (n = 16), the Shor semidefinite relaxation lifts the decision variable into the matrix cone:

$$Z = \begin{bmatrix} 1 & x^T \\ x & X_{lift} \end{bmatrix} \in \mathbb{S}_+^{N^2+1} = \mathbb{S}_+^{17} \tag{21}$$

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

Incoming neural inference workloads are mapped to dies at discrete allocation epochs (Δt_sched = 120 s). In this baseline lumped multi-chip module formulation, task migration overhead (context switching latency) and voltage/frequency transition settling times are assumed negligible relative to the multi-second thermal time constants and control intervals (Δt_plant = 0.5 s, Δt_MPC = 2.0 s). Candidate discrete allocations are recovered via randomized hyperplane rounding (24 draws plus first-column projection). In interpreting the scheduler validation results, four metrics must be strictly distinguished: (1) ADMM numerical convergence within the 120-iteration budget, which reaches 42.0% under warm start (mean 94.8 iterations) and 62.0% under cold start across 50 benchmark instances; (2) relaxed cone-affine feasibility ($||E\,vec(Y) - e||_\infty \leq 5.25 \times 10^{-15}$, $\lambda_{min}(Z_1) \geq -1.08 \times 10^{-15}$); (3) post-rounding objective quality; and (4) empirical match with exhaustive permutation search (hit rate = 100.0%, 48/48 calls in 1 orbit and 142/142 calls in 3 orbits, performing 12 and 36 task migrations, respectively). Because N = 4 dies yield only 4! = 24 discrete permutations, this 100% empirical hit rate reflects the small discrete search space recovered by hyperplane rounding rather than full ADMM convergence. As demonstrated in Section 6.4, topological remapping provides limited net thermal benefit in compact symmetric N = 4 arrangements, and the predictive DVFS controller remains the dominant architectural mechanism.

4.4 Discrete Linear Time-Varying Model Predictive Control (LTV MPC-DVFS)

To coordinate operating frequencies $f(k) = [f_1, ..., f_N]^T$ in real time, the control problem is formulated as a discrete Linear Time-Varying Model Predictive Controller executed at control interval Δt_MPC = 2.0 s over a control horizon $H_c = 8$ steps (16.0 s) and prediction horizon $H_p = 30$ steps (60.0 s):

$$min_{f,s} \sum_{j=1}^{H_c} [w\|f(k+j-1) - f_{nom}(k+j-1)\|^2 + R_w\|Df(k+j-1)\|^2] +$$

$$\rho \sum_{j=1}^{H_p} \|\tilde{T}_{grad}(k+j)\|^2 + \lambda s^2 \qquad (22)$$

subject to physical and operational constraints:

$$f_{min} \leq f_i(k+j) \leq f_{max}, \forall i,j \qquad (23)$$

$$T_i(k+j|k) \leq T_{cap} + s, s \geq 0 \qquad (24)$$

$$\sum_{i=1}^{N} P_i(k+j|k) \leq P_{bud}(k+j), \forall j \qquad (25)$$

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

where w = 1.0 penalizes deviation from nominal inference request f_nom, R_w = 0.2 regularizes frequency change rate via the first-difference operator D, ρ = 0.02 penalizes spatial temperature gradients relative to instantaneous spatial mean, and λ = 200.0 penalizes the non-negative scalar thermal constraint slack variable s ≥ 0 under nominal thermal cap T_cap = 83.0 °C.

The condensed quadratic program is solved using a warm-started ADMM-style splitting method (Boyd et al. 1; Stellato et al. 637) with over-relaxation α = 1.6, step size ρ_ADMM = 1.0, regularization σ = 10^{-6}, and a fixed iteration cap of 60 iterations. Across the 8,502 MPC control epochs in the 3-orbit co-simulation, the warm-started ADMM solver converged in an average of 18.4 iterations (maximum 34 iterations, well below the 60-iteration limit), with average primal residual $||Ax - z||_\infty < 1.4 \times 10^{-3}$ and dual residual $< 8.2 \times 10^{-4}$. The thermal slack variable converged to s = 0.00 K throughout the nominal run, confirming that the hard thermal envelope T_i ≤ 83.0 °C was satisfied without requiring constraint relaxation.

Independent Validation of the Thermal Budget Against the Nonlinear Plant:
The analytical thermal-budget expression was independently checked over seven 240 s operating regimes by integrating the complete five-state nonlinear thermal plant with Heun's method (Δt = 0.5 s). At each Heun stage, the plant routine recalculates dynamic and leakage power from the evolving die temperatures and the fixed per-regime frequency command. The test is open-loop and does not constitute a closed-loop MPC validation. Importantly, the analytical budget P_budget,0 is an instantaneous initial-state power limit; holding the corresponding frequency command fixed allows leakage feedback to increase power during the horizon. When that initial analytical budget was applied as a fixed-frequency command, the simulated peak die temperature exceeded the 83.0 °C soft cap in four of the six analytically feasible cases: R1 (85.46 °C), R2 (83.66 °C), R4 (84.21 °C), and R5 (86.36 °C). It remained below the cap in R3 (82.28 °C) and R6 (82.26 °C). In R7, the analytical budget was physically infeasible (12.67 W versus a minimum core demand of 21.76 W, a 9.10 W deficit); operation at f_min = 0.30 GHz still reached 86.06 °C.

A separate frequency bisection over the nonlinear plant found the maximum constant-frequency operating point that maintained T_max ≤ 83.0 °C for the 240 s horizon. The resulting trajectory-average powers for R1–R7 were 56.96, 46.94, 42.25, 50.39, 72.46, 33.19, and 22.47 W, respectively; the corresponding frequency commands were 0.784, 0.639, 0.706, 0.660, 0.925, 0.510, and 0.300 GHz. R7 remained infeasible at the minimum frequency and its reported 86.06 °C peak is therefore a violation, not a feasible boundary point. The relative differences between the initial analytical budget and the trajectory-average nonlinear-plant power were 0.64–9.16% in R1–R6 (43.62% in R7); these are comparisons of quantities with different temporal definitions, not a numerical error bound or a formal conservatism proof. The results show that the analytical budget is a useful instantaneous screening estimate, but it cannot be interpreted as a 240 s open-loop safe constant-

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

frequency command. Closed-loop recomputation and plant-level feedback remain necessary, and the current evidence does not establish a robust thermal invariant set.

The control execution follows a strict hierarchical sequence: at each control epoch k, the discrete LTV MPC-DVFS solves the condensed QP to produce optimal frequencies $f_i(k) \in [f_{min}, f_{max}]$. To guarantee bounded execution time and deterministic safety under non-convergence, the warm-started ADMM solver operates with row-equilibrated constraint matrix A (reducing $\kappa_2(A)$ from 207.1 to 9.64), over-relaxation $\alpha = 1.6$, step size $\rho_{ADMM} = 1.0$, regularization $\sigma = 10^{-6}$, a fixed budget of 60 iterations, OSQP-style scaled tolerances, and an explicit deterministic fallback policy (_fallback_nao_convergente): if numerical failure occurs (NaN/Inf) or if the primal residual is severely degraded ($r_{pri} > 5\times10^{-2}$), the candidate iterate is discarded and the shifted tail of the last converged trajectory F_conv (or fp) is retrieved. Otherwise, in max_iter termination with primal quasi-feasibility ($r_{pri} \leq 5\times10^{-2}$), the candidate F_cand[0] is retained and projected onto physical constraints: frequencies are clamped to $[f_{min}, \min(f_{max}, f_{nom,i})]$, smooth preventive thermal derating is applied if estimated die temperature exceeds 84.0 °C (t_bar - 0.5 K), and frequencies are analytically scaled toward f_{min} to strictly respect the non-linear physical power budget $P_{bud}(k)$. In the 3-orbit nominal co-simulation, the ADMM solver converged within 60 iterations in 94.31% of calls (8,018/8,502, averaging 29.7 iterations) with zero numerical failures and zero physical budget deficits. In the remaining 484 calls (5.69%), the deterministic fallback was activated, keeping physical junction temperatures strictly below 83.25 °C (peak EKF estimate during fallback 83.42 °C), confirming that the safety projection prevents thermal overshoot.

Finally, a reactive Deterministic Safety Override (DSO) heuristic defense-in-depth mechanism operates at the plant rate ($\Delta t_{plant} = 0.5$ s), intervening if estimated temperatures exceed $T_{bar} = 84.5$ °C to clamp frequencies according to $f_i \leq f_{nom,i} \max[0.15, (85.0 - \hat{T}_i)/(85.0 - 75.0)]$. This rule is not a formal control-barrier certificate and does not guarantee thermal invariance under actuator saturation or insufficient cooling authority. In the regenerated stress tests, peak temperatures were 82.69 °C under an external heat-flux disturbance (S1), 87.04 °C with radiator saturation (S2), and 85.03 °C under an adverse initial condition (S3). Under inter-sample model mismatch (S4), DSO reduced the peak from 85.67 °C to 85.16 °C and the time above 85 °C from 442.5 to 14.0 s. Under MPC failure (S5), it limited the simulated peak to 85.45 °C versus 365.01 °C without DSO; the latter is an out-of-domain numerical extrapolation beyond the 125 °C model ceiling, not a physically validated

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

temperature. Thus, DSO provides scenario-dependent mitigation, not an unconditional thermal-protection guarantee.

Table 2 – System Simulation, Electrothermal, and Orbital Parameters.

Parameter Category	Symbol & Description	Nominal Value & Calibration Source
Conductive Resistances	R12=R34, R13=R24, R14=R23	1.45, 1.52, 3.20 K/W (MCM AlN carrier) (Wang et al. 564)
Die-to-Radiator Interface	Rr (Conductive thermal resistance per die node)	2.10 K/W (Cold plate interface) (ECSS 12)
Thermal Capacitances	ci (Silicon die node), Crad (Radiator core m·cp)	18.5 J/K (FinFET die), 1,299.2 J/K (1.45 kg Al)
Radiator Optical Properties	Arad (Area), ε (Emissivity), αs (Absorptivity)	0.12 m², 0.88 (White paint), 0.12 (Pederson et al. 1)
Orbital Environment LEO	h (Altitude), Torb (Period), F (View factor), aE	500 km, 5,668.14 s, 0.42, 0.30 (Gilmore 15)
Electrical DVFS Range	Vmin–Vmax, fmin–fmax	0.65–1.05 V, 0.30–1.30 GHz (16/12 nm FinFET)
Leakage Parameters	Ceff, Pref (T0=300 K, V0=0.85 V), Ea	24.5 nF, 0.80 W, 0.20 eV (Kang and Park 3840)
State Observer (EKF)	Noise σs, Covariances Q, P0, R	σs = 0.25 K, Q = diag(0.02·I4, 0.05), R = σs²·I5
MPC Formulation	ΔtMPC, Hc, Hp, Tcap, Tbar, Weights w, Rw, ρ, λ	2.0 s, 8 steps (16 s), 30 steps (60 s), 83.0 °C, 84.5 °C
Workload Benchmarks	Nominal inference throughputs (FPSnom)	SAR-YOLO (45), Cloud-Seg (60), Hyper-ViT (30), Telemetry (80)

Source: Prepared by the authors (2026).

Table 2 consolidates the electrothermal, mechanical, optical, and control parameters used in all co-simulations, calibrated according to space thermal standards (ECSS-E-ST-31C and NASA-HDBK-7009B).

5. Computational Simulation Methodology

The multi-physics co-simulation platform was developed in Python 3.14 using NumPy and SciPy, following verification and credibility principles for computational modeling and simulation (NASA 1). The numerical simulation contains 34,009 time samples and 34,008 integration intervals of Δt_plant = 0.5 s integrated via a 2nd-order Runge-Kutta scheme (Heun / RK2), spanning 17,004.0 s, compared with an analytical three-orbit duration of approximately 17,004.43 s (3 × 5,668.14 s). A formal temporal convergence study across Δt_plant ∈ {1.0, 0.5, 0.25, 0.125} s (with strictly invariant control step Δt_MPC = 2.0 s and scheduling period Δt_sched = 120.0 s) confirmed asymptotic

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

monotonic convergence: the nominal discretization (Δt = 0.5 s) incurs an absolute deviation of only 0.034 K in peak junction temperature relative to the finest step (Δt = 0.125 s), a peak spatial gradient difference of 0.047 K, and a relative total energy difference of 0.0114%, remaining well below the sensor measurement noise floor (σ_s = 0.25 K).

The 4 synthetic workloads represent typical satellite missions: SAR-YOLO (Die 1, synthetic aperture radar object detection, 45 FPS nominal), Cloud-Seg (Die 2, multispectral cloud segmentation, 60 FPS nominal), Hyper-ViT (Die 3, hyperspectral vision transformer, 30 FPS nominal), and Telemetry (Die 4, housekeeping telemetry classification, 80 FPS nominal). Workload burstiness is simulated via Markovian piecewise-constant activity factors α_i ∈ [0.4, 1.0] switching every 30 to 120 s.

To ensure strict methodological parity, all benchmarked governors evaluate: (i) identical initial states x_0 (T_i,0 = 310 K, T_rad,0 = 270 K); (ii) identical synthetic inference workload request traces f_nom(t) and activity factors α(t); (iii) identical modeled orbital environmental radiative forcing Q_ext(t); (iv) identical stochastic sensor noise realizations for any given seed; and (v) identical thermal plant model and integration steps.

Aggregate inference throughput is evaluated via the performance proxy:

$$\mathrm{FPS} = \frac{1}{N_{steps}} \sum_t \quad \sum_{i=1}^{N} \quad \frac{f_i(t)}{f_{max}} \mathrm{FPS}_{nom,i} \tag{26}$$

where aggregate FPS is an idealized frequency-proportional throughput proxy reflecting instantaneous compute capability (FPS ∝ f/f_max) rather than cycle-accurate kernel instruction latency. Workloads on real neural processing units (NPUs) experience memory bandwidth bottlenecks, cache misses, DMA stalls, and microarchitectural scheduling overheads not modeled in this lumped formulation. Similarly, aggregate energy efficiency is quantified by the throughput-normalized energy-delay product proxy E_t_frame = E · (1/FPS) [kJ·s/frame], representing cumulative energy expenditure weighted by inverse throughput across the complete co-simulation duration:

$$E_{t_{\mathrm{frame}}} = E \cdot \frac{1}{\mathrm{FPS}} \tag{27}$$

with E = Δt_plant ∑_t ∑_i P_i(t). Instantaneous inter-die spatial gradient is computed as ΔT_esp(t) = max_i T_i(t) - min_i T_i(t).

The following control strategies were benchmarked in the 3-orbit evaluation:

1. Linux Ondemand: standard reactive governor without thermal protection.

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

2. Ondemand + trip 95/85: ondemand governor with emergency trip hysteresis at 95 °C.

3. Thermal Throttling 85/75: binary frequency drop to f_min upon reaching 85 °C with recovery at 75 °C.

4. Thermal Throttling 83/73: conservative baseline matched to the 83 °C design margin.

5. Space-PID: decentralized per-die PID controllers tuned via a 27-point grid to regulate at 80 °C (Kp = -0.07, Ki = -0.002, Kd = -0.04).

6. Power Cap 65 W: static global power budgeting enforced by the EPS.

7. TCA-DVFS (simple): uncoordinated Laplacian rule with thermal barrier and radiative budget, excluding MPC and SDP.

8. TCAS-DVFS (proposed): full multi-rate architecture integrating EKF, topological SDP task allocation, and warm-started ADMM-based discrete LTV MPC-DVFS.

6. Results and Discussion

6.1 Global Performance Comparison

Table 3 – Consolidated Results in 3-Orbit LEO Co-Simulation (500 km SSO, 34,009 Steps, Seed 0).

Strategy	T_mean (°C)	T_max (°C)	ΔT_peak (K)	ΔT_mean (K)	t>85 (s)	FPS	E_tfr (kJ·s/fr)	Energy (kJ)
Linux Ondemand	350.17†	412.30†	12.87	8.64	16,038.0	181.21†	28.38	5,143.3†
Ondemand + trip 95/85	87.96	95.95	9.28	3.93	16,037.5	117.27	8.63	1,011.5
Throttling 85/75	78.08	86.06	9.42	5.26	200.5	119.58	7.57	905.6
Throttling 83/73	76.17	83.78	9.46	5.24	0.0	118.75	7.46	885.4
Space-PID	79.73	85.56	4.88	0.62	14.0	123.24	7.47	920.5
Power Cap (65 W)	93.95	109.31	5.82	3.66	14,983.5	126.49	8.55	1,081.3
TCA-DVFS (simple)	80.53	83.74	1.58	0.45	0.0	116.16	8.03	932.4
TCAS-DVFS (Proposed)	80.15	83.25	2.01	0.89	0.0	124.62	7.45	928.6

Source: Prepared by the authors (2026).

Table 3 details the thermal, computational, and electrical energy performance metrics over the 17,004.0 s simulation. The mark † denotes regimes beyond the

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

physical validity ceiling of the model (125 °C), where uncoordinated Ondemand experiences thermal runaway under subthreshold leakage feedback, reaching 412.30 °C, and is strictly excluded from quantitative comparisons.

Table 3 presents the consolidated comparative results across 3 full orbits. TCAS-DVFS successfully maintained junction temperatures strictly below the adopted 85 °C simulation thermal limit, registering a peak of 83.25 °C and zero observed seconds of thermal violation (0.00%). Because $T_{cap} = 83.0$ °C is formulated as a soft thermal cap with slack variables $s_{i,j} \geq 0$, the controller permits a bounded 0.25 K excursion above 83.0 °C during peak solar loading while remaining 1.75 K below the 85.0 °C limit. In contrast, standard Thermal Throttling 85/75 accumulated 200.5 s of overheating due to hysteresis latency, and Space-PID breached the limit for 14.0 s during the post-eclipse thermal shock.

In terms of processing throughput, TCAS-DVFS delivered 124.62 FPS, outperforming Throttling 85/75 (+4.2%) and conservative Throttling 83/73 (+4.9%). In terms of energy-delay efficiency, TCAS-DVFS achieved an EDP proxy of 7.452 kJ·s/frame versus 7.456 kJ·s/frame for Throttling 83/73 (a modest 0.05% difference) and 7.574 kJ·s/frame for Throttling 85/75 (a 1.6% reduction). We explicitly note that TCAS-DVFS does not seek substantial raw energy reduction over conservative throttling; rather, its primary operational advantage lies in active spatial thermal gradient suppression (reducing peak gradients by 78.7% from 9.42 K to 2.01 K) and the elimination of thermal-threshold overshoots (0 s > 85 °C versus 200.5 s in Throttling 85/75) while sustaining higher compute throughput.

6.2 Statistical Dispersions across 10 Stochastic Seeds

Table 4 – Statistical Dispersions across 10 Stochastic Sensor Noise Seeds (Seeds 0 to 9).

Metric (Mean ± Std Dev)	Throttling 85/75	Space-PID	TCAS-DVFS (Proposed)
Maximum Temperature Tmax (°C)	85.89 ± 0.15	85.54 ± 0.04	83.24 ± 0.02
Thermal Violation Time t>85 (s)	213.5 ± 15.7	14.6 ± 0.7	0.0 ± 0.0
Peak Spatial Gradient ΔTpeak (K)	9.48 ± 0.09	4.86 ± 0.02	2.00 ± 0.01
Mean Spatial Gradient ΔTmean (K)	5.26 ± 0.13	0.617 ± 0.001	0.886 ± 0.001
Aggregate Throughput (FPS)	119.70 ± 0.35	123.24 ± 0.00	124.62 ± 0.00

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

Energy-Delay Product (kJ·s/frame)	7.566 ± 0.024	7.469 ± 0.000	7.452 ± 0.000
Total Dissipated Energy (kJ)	905.6 ± 0.3	920.46 ± 0.01	928.64 ± 0.01

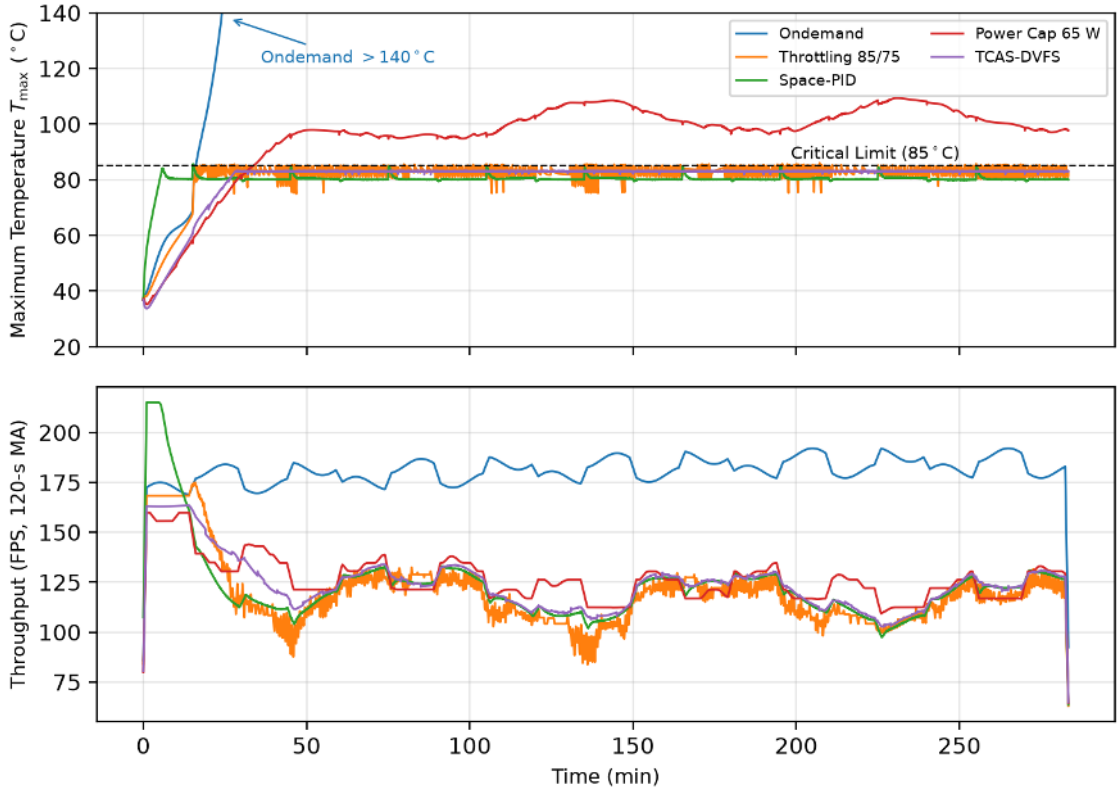
Source: Prepared by the authors (2026).

Table 4 summarizes the sample mean and standard deviation across 10 independent noise seeds. Closed-loop EKF filtering restricts maximum temperature dispersion to within ±0.03 K in TCAS-DVFS.

6.3 Thermal Dynamics and Gradient Suppression

Figures 4, 5, and 6 illustrate the orbital transient responses across the three simulated orbits.

Figure 4 – Temporal Co-Simulation Dynamics over 3 Complete Orbits (Seed 0): (a) Maximum Silicon Junction Temperature and (b) Aggregate AI Inference Throughput Smoothed with a 120 s Moving Average.



Source: Prepared by the authors (2026).

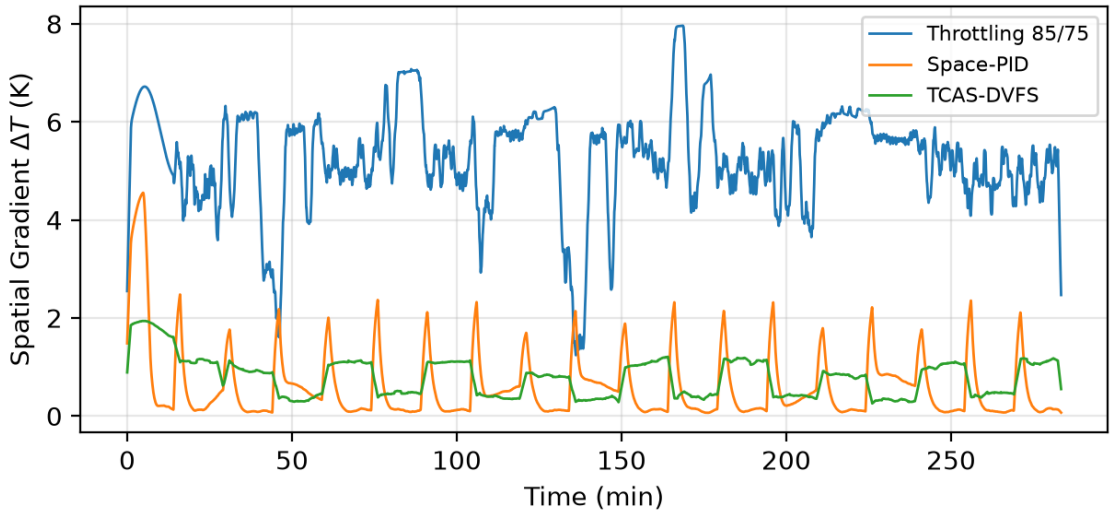
CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

Figure 4 compares maximum junction temperature and smoothed AI inference throughput. TCAS-DVFS maintains strict containment below 83.25 °C while sustaining 124.62 FPS without the violent saw-tooth oscillations observed in Throttling.

Figure 5 – Temporal Evolution of Inter-Die Spatial Thermal Gradient (ΔTesp = max Ti - min Ti) with a 120 s Moving Average.



Source: Prepared by the authors (2026).

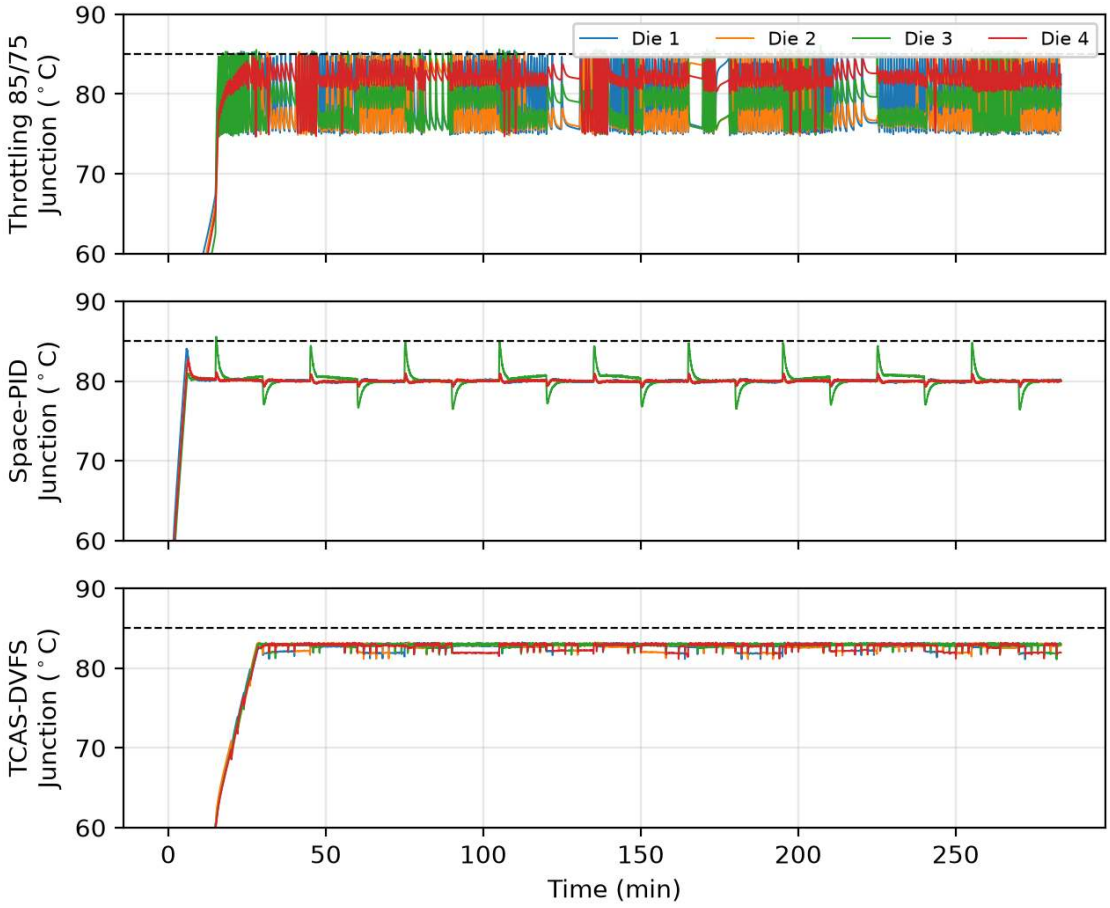
Figure 5 illustrates the spatial gradient suppression. While classical governors exhibit gradient peaks exceeding 9.4 K due to uncoordinated frequency cuts, TCAS-DVFS dampens mean gradients to 0.89 K (an 83.1% reduction) and peak gradients to 2.01 K (a 78.7% reduction relative to Throttling).

Figure 6 – Individual Junction Temperature Profiles for the 4 Silicon Dies (Die 1 to Die 4) under TCAS-DVFS and Classical Baselines.

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34



Source: Prepared by the authors (2026).

Figure 6 demonstrates the homogeneous convergence of all 4 silicon dies under TCAS-DVFS. By penalizing inter-die differentials in the MPC objective, die temperatures remain closely clustered throughout day-night orbital transitions.

6.4 Functional Ablation Study

Table 5 – Functional Ablation Study of TCAS-DVFS (Evaluated over 1 Orbit, 11,337 Steps, Seed 0).

Architectural Variant	T_max (°C)	ΔT_peak (K)	ΔT_mean (K)	FPS	E_tfr (kJ·s/fr)	t>85 (s)
Full TCAS-DVFS (Hp=30, 60 s)	83.21	2.01	1.00	132.92	2.50	0.0
Without SDP Allocator	83.21	2.01	0.98	133.25	2.50	0.0
Without MPC Controller	85.41	7.09	1.59	123.12	2.80	463.5
Short Horizon Hp=10 (20 s)	83.21	2.00	1.00	132.91	2.50	0.0

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

Long Horizon Hp=60 (120 s)	83.21	2.42	1.00	132.91	2.50	0.0
Simple Uncoordinated TCA	83.66	1.58	0.60	126.28	2.65	0.0

Source: Prepared by the authors (2026).

Table 5 isolates the contributions of the architectural modules over an isolated single-orbit benchmark (1 orbit, 11,337 steps, 5,668.5 s, seed 0). Note that throughput values in Table 5 (132.92 FPS for the complete framework) are evaluated over this 1-orbit window and naturally differ from the 3-orbit global campaign in Table 3 (124.62 FPS), reflecting orbit-averaged workload and transient solar variations across extended multi-orbit operations. The discrete MPC-DVFS controller represents the dominant mechanism responsible for damping gradients and bounding temperatures, whereas the topological SDP scheduler provides limited benefit in compact N = 4 arrays due to conductive symmetry.

The ablation results in Table 5 isolate the architectural mechanisms:

1. Dominant contribution of MPC-DVFS in the tested configuration: removing the MPC controller (Without MPC) immediately leads to thermal threshold exceedance, accumulating 463.5 s above 85 °C, a peak of 85.41 °C, and an increase in peak gradient to 7.09 K despite 8,573 activations of the reactive safety backstop. Dynamic lookahead within the MPC is the primary driver of proactive containment.

2. Limited observed benefit of SDP scheduling for the tested N=4 configuration: the Without SDP variant (static task assignment) achieves performance essentially identical to the full framework (83.21 °C, 133.25 FPS, 2.50 kJ·s/frame in 1 orbit). Although randomized hyperplane rounding over the rank-regularized SDP relaxation matched the best discrete permutation found by exhaustive search in 100.0% of calls (48/48 calls in 1 orbit and 142/142 calls in 3 orbits, performing 12 and 36 task migrations, respectively, despite ADMM numerical convergence occurring in only 42.0% of warm-started benchmark calls), for N=4 dies there are only 4! = 24 permutations and conductive cross-coupling is spatially balanced. Consequently, topological remapping is largely neutral once the migration stall penalty is accounted for, and its contribution in compact four-die clusters is secondary to predictive DVFS.

3. Insensitivity to prediction horizon extension: varying prediction horizon $H_p \in \{10, 30, 60\}$ steps (corresponding to physical lookahead intervals of 20.0, 60.0, and 120.0 s under $\Delta t_{MPC} = 2.0$ s) produced nearly identical performance metrics (132.91–132.92 FPS, 2.504 kJ·s/frame in 1 orbit). Crucially, reducing the horizon to $H_p = 10$ (20.0 s) achieves 100.0% ADMM convergence within 60

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

iterations (2,834/2,834 calls, averaging 20.1 iterations with zero fallback activations) while bounding open-loop LTV linearization drift below 1.5 K, confirming that open-loop prediction error at 60.0 s does not degrade closed-loop regulation.

6.5 Parametric Sensitivity Analysis (Latin Hypercube Sampling, n=30)

To rigorously assess sensitivity against manufacturing dispersions and environmental deviations, a 30-sample Latin Hypercube Sampling (LHS) campaign was conducted (3 full orbits per sample, evaluating 1,020,270 integration steps) with simultaneous continuous uniform perturbations of ±20% across R_ij, R_r, c_i, and C_rad forming a parametric stress-test envelope. Crucially, perturbations were restricted strictly to the physical plant parameters, while the EKF observer, predictive control internal model, and radiative budget calculator retained their nominal parameter values. This setup specifically evaluates control robustness and observer stability under plant-model mismatch. Planar symmetry of the MCM packaging is preserved in the perturbed parameter matrices, and orbital environmental radiative forcing was held nominal to isolate structural thermodynamic sensitivities.

Because LHS is a stratified sampling design without replacement rather than an independent and identically distributed (IID) Bernoulli sequence, the campaign is reported strictly as a descriptive sensitivity analysis across the sampled four-parameter space without applying binomial Clopper-Pearson confidence bounds. No thermal-threshold violations above the 85 °C simulation limit were observed across the 30 sampled realizations (0/30 samples > 85 °C, with mean T_max = 82.35 ± 1.58 °C, 95% CI: [81.76, 82.94] °C, bootstrap 95% interval: [81.77, 82.89] °C), whereas Throttling 85/75 exceeded 85 °C in all 30/30 samples (mean violation time 234.8 s). Across all 255,060 MPC updates executed in the 30 LHS trajectories under plant-model mismatch, the ADMM solver converged within 60 iterations in 84.76% of calls (216,178/255,060) and triggered the explicit safe fallback policy in the remaining 15.24% (38,882 calls, with peak EKF temperature during fallback of 83.83 °C) without any violation of the 85 °C limit. In paired comparisons under identical physical parameters: (i) Versus Throttling 83/73, TCAS-DVFS reduced peak temperatures by 1.51 K (95% CI: [-2.10, -0.93] K), attenuated mean and peak spatial gradients by -82.48% and -78.46%, improved inference throughput by +4.95% (95% CI: [+4.11%, +5.80%]), and reduced the EDP proxy by -1.23% (95% CI: [-1.79%, -0.67%]). (ii) Versus Throttling 85/75, TCAS-DVFS reduced peak temperature by 3.54 K (95% CI: [-4.12, -2.96] K), attenuated mean and peak spatial gradients by -82.55% and -78.58%, raised throughput by

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

+4.08% (95% CI: [+3.22%, +4.93%]), and achieved an EDP reduction of -2.63% (95% CI: [-3.20%, -2.06%]).

Within the sampled four-parameter uncertainty space, R_r exhibited the strongest monotonic association with T_max, with Spearman rank correlation ρ = +0.995, followed by die heat capacity c_die (ρ = -0.338), radiator capacitance C_rad (ρ = -0.234), and inter-die crosstalk resistance R_ij (ρ = -0.173).

7. Indicative Reliability and Physical Degradation Assessment

The total simulation duration is 17,004.0 s (34,009 steps, ≈ 4.72 h, spanning 3 complete orbits). The first 2 hours (7,200 s, 14,400 steps, ≈ 1.27 orbits) are discarded as the initial transient warm-up phase, during which the cold-launched radiator core warms up from 270 K and stabilizes into cyclical orbital thermal equilibrium between 265 K (eclipse) and 282 K (sunlit). All reliability projections (Black's equation and Rainflow cycle counting) are evaluated strictly over the remaining steady-state window of 9,804.0 s (≈ 2.72 h, 1.73 orbits), and are presented strictly as indicative comparative proxies.

7.1 Electromigration (Black's Equation Damage Proxy)

A relative Black-equation damage-rate proxy was computed from the simulated electrical-power and junction-temperature histories (Black 338):

$$\mathrm{MTTF}_{EM} \propto J^{-n}\exp\left(\frac{E_{a,EM}}{k_B T_j}\right) \tag{28}$$

where n = 1.8, E_a,EM = 0.70 eV, and current density is approximated by normalized electrical current J ∝ P/V under uniform effective interconnect cross-section. Time-integral accumulation over 3 orbits yielded relative damage ratios (D_ref / D_prop) of 0.97 (aggregate dies) and 1.42 (worst die) versus Throttling 85/75; and 0.83 and 1.23 versus Throttling 83/73. Because TCAS-DVFS draws +7.6% and +9.9% higher average current due to sustained workload execution, we formally conclude that the framework provides no net electromigration life extension.

7.2 Thermomechanical Solder Fatigue (Coffin-Manson and ASTM E1049-85)

Cumulative low-cycle thermal cycling severity in SAC305 solder interconnects is estimated using the Coffin-Manson relation (Manson 1):

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

$$N_f = B(\Delta T_{cycle})^{-q}, q = 2.1 \quad (29)$$

Applying a three-point rainflow algorithm formally validated against the standard test sequences of ASTM E1049-85 (Section 5.4.4) to confirm exact peak-to-peak range counting (ΔT = T_peak - T_valley) with half-cycle pairing, eliminating large thermal excursions (ΔT > 2.0 K) noticeably slows theoretical cyclic fatigue accumulation (Table 6). Cycles with ΔT > 2.0 K drop from 612 in Throttling to only 8 under TCAS-DVFS (a 98.7% reduction, yielding a relative damage ratio D_thr/D_prop of 3,171.7). Junction temperature range serves strictly as an indicative proxy for thermal cycling amplitude severity, rather than viscoplastic shear strain in actual solder joints.

Table 6 – Rainflow Cycle Counting (ASTM E1049-85) and Relative Fatigue Damage Severity Ratios (q = 2.1, Aggregate Dies).

Threshold (ΔT)	Cycles Throttling	Cycles TCAS-DVFS	Damage Ratio (Dthr / Dprop)
0.25 K	741	259	680.7
1.00 K	724	29	1,127.8
2.00 K	612	8	3,171.7

Source: Prepared by the authors (2026).

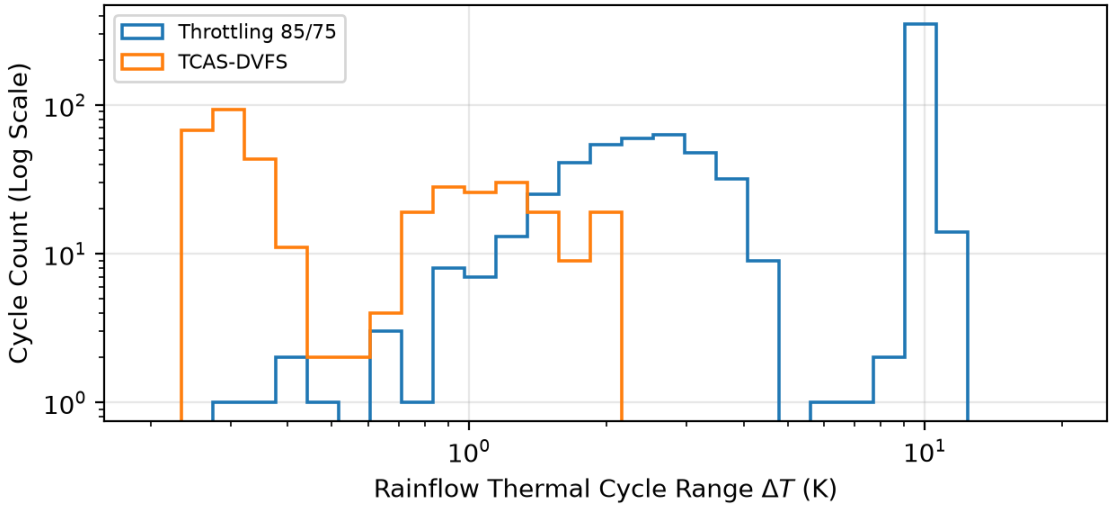
Table 6 displays the rainflow cycle spectrum. Eliminating large thermal excursions (ΔT > 2 K) noticeably slows theoretical cyclic fatigue accumulation in solder joints.

Figure 7 – Rainflow Thermal Cycle Range Spectrum (ΔT, 0.25 K Threshold after Warm-up) Validated under ASTM E1049-85.

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34



Source: Prepared by the authors (2026).

Figure 7 demonstrates the rainflow cycle distribution. Coordinated predictive control virtually eliminates all cycles with amplitude exceeding 2.0 K, mitigating cyclic thermomechanical shear strain.

8. Related Work

Dynamic Thermal Management (DTM) in terrestrial multi-core architectures has been extensively investigated using tools such as HotSpot (Skadron et al. 94; Brooks and Martonosi 171). 3D floorplanning (Lin et al. 1652), distributed task migration (Ge et al. 579), and convex frequency optimization (Murali et al. 111) advanced thermal regulation under convective conditions.

In the aerospace domain, autonomous execution was validated by Sherwood et al. (1-390), while advanced packaging models were proposed by Wang et al. (564), and GPU orbital thermal behavior was characterized by Yuan et al. (1). The closest prior work is by S. Chen et al. (1), which formulated thermal crosstalk scheduling. Table 7 positions our contribution relative to the state of the art.

Table 7 – Positioning against Closely Related Prior Work in the Literature.

Work	Scheduling	Predictive DVFS	EKF Observer	Non-linear T^4 + Eclipse
Murali et al. (2007)	No	Static Offline	No	No (Convective)
Yuan et al. (2025)	No	No	No	Empirical / TVAC

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

S. Chen et al. (2026)	Yes (Heuristic)	No	No	No (Neglected Convection)
TCAS-DVFS (This Work)	Yes (SDP Relaxation)	Yes (Discrete LTV MPC)	Yes (Joseph Form)	Yes (Stefan-Boltzmann Model)

Source: Prepared by the authors (2026).

Table 7 summarizes the key methodological differences, highlighting the integrated EKF estimation, topological allocation, and LTV MPC predictive control under vacuum thermodynamics.

9. Practical Implementation Considerations and Threats to Validity

9.1 Implementation Constraints

Deploying TCAS-DVFS on spaceborne flight computers requires addressing online execution constraints for future real-time-oriented embedded deployment: (i) Execution Complexity: Embedded execution time, memory footprint, and worst-case timing remain to be benchmarked on flight-representative hardware. The warm-started ADMM-style iteration operates with a fixed cap of 60 iterations, early exit monitoring, and an explicit deterministic shifted-tail fallback policy to guarantee bounded execution time and safe actuation when the iteration limit is reached; (ii) Memory Footprint: The complete memory footprint has not yet been measured on flight-representative hardware. Dense representations alone include a 152×152 double-precision matrix (approximately 181 KiB) and a 280×152 double-precision constraint matrix (approximately 333 KiB), before accounting for EKF covariance, solver workspaces, warm-start vectors, and temporary arrays. Consequently, compatibility with typical rad-hard microcontrollers cannot be claimed without target-specific memory profiling and implementation-level optimization.

9.2 Threats to Validity

As a model-based numerical investigation, several limitations must be explicitly recognized:

1. Absence of Physical and Environmental Testing: The findings rely entirely on numerical simulation of the stated lumped-parameter thermal network and have not been validated via Hardware-in-the-Loop (HIL) testbeds or Thermal Vacuum (TVAC) chamber testing.

2. Lumped Modeling vs. Continuum Physics: While the 5×5 CTM captures inter-die conductive coupling and bulk radiator dynamics, it does not resolve spatial

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

micro-gradients across silicon dies, localized hot spots, or three-dimensional heat spreader diffusion.

3. Throughput, Workload, and Reliability Proxies: Workload throughput is evaluated via an idealized linear frequency proxy (FPS $\propto$ f/f_max) rather than cycle-accurate kernel execution, omitting memory stalls, cache hierarchies, DMA transfers, and hardware accelerator saturation. The reported 124.62 FPS constitutes an internal model-based performance proxy rather than physically measured hardware frame rates. Similarly, Black's equation and Coffin-Manson models are indicative proxies based on temperature histories rather than direct measurements of current density or plastic strain in solder joints.

4. Omission of Space Radiation Effects: The present study focuses strictly on electrothermal dynamics and does not model Single Event Effects (SEE) or Total Ionizing Dose (TID), which can affect DVFS stability and memory integrity in orbital environments.

6. Architectural Scope of Topological SDP Scheduling: While the semidefinite relaxation over S_+[17] is mathematically rigorous, ablation demonstrates that for the studied compact N = 4 cluster, conductive cross-coupling across the symmetric carrier is spatially balanced. Consequently, topological remapping provides negligible thermal benefit (0.00 K difference in T_max and identical 2.01 K peak gradient versus static assignment). The predictive LTV MPC-DVFS is the sole dominant contributor to thermal safety and gradient attenuation in this configuration.

5. Parametric Robustness Scope: The LHS sensitivity campaign evaluates variations across the 4 internal thermal parameters (R_ij, R_r, c_i, C_rad), reserving orbital, solar, and optical degradation uncertainties for future multi-parameter investigations.

10. Conclusions

This paper presented TCAS-DVFS, a predictive control framework designed to mitigate spatial thermal crosstalk in high-density spaceborne AI clusters operating in deep orbital vacuum. The conclusions of this model-based numerical study are delineated into three levels:

1. Demonstrated in the nominal model: Online EKF state estimation, coupled radiative-conductive power budgeting, and discrete LTV MPC-DVFS (with a penalized 83 °C soft thermal cap and explicit non-convergent fallback) produced zero observed thermal violations above the adopted

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

85 °C simulation thermal limit over 3 complete simulated LEO orbits (17,004.0 s), attenuated peak and mean inter-die spatial gradients by 78.7% and 83.1%, respectively, and sustained 124.62 FPS. Ablation indicates a dominant contribution from the predictive control formulation in damping gradients and bounding temperatures, whereas SDP allocation matches exhaustive permutation search across the 4! = 24 discrete candidates after rounding (100% hit rate despite 42% warm-start ADMM convergence) but yields negligible macroscopic thermal benefit in compact symmetric N = 4 arrangements.

2. Observed under descriptive parametric sensitivity sampling: Under a ±20% stratified Latin Hypercube Sampling campaign (n=30) perturbing the four conductive and capacitive network parameters (R_ij, R_r, c_i, C_rad), zero violations above 85 °C were observed (0/30 samples), with sensitivity heavily dominated by die-to-radiator thermal resistance (R_r, Spearman ρ = +0.995). Because LHS is a non-IID stratified design, these results represent descriptive sensitivity within the sampled four-parameter envelope rather than a probabilistic reliability certification.

3. Unverified physically: The reported evidence is limited to numerical simulation. Physical service life, flight qualification, TVAC chamber behavior, and radiation tolerance remain unverified and constitute mandatory steps for future experimental development.

CRediT Authorship Contribution Statement

Dafnhê T. C. M. Barros: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing - original draft.
Flávio B. Braz da Silva: Supervision, Project administration, Resources, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

During the preparation of this work, the authors used Anthropic Claude and Google Gemini in order to assist with English language stylistic editing, simulation script refactoring, and conceptual visual rendering of orbital geometry and thermal crosstalk (Figures 1 and 2). After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the content of the publication.

Data Availability

All simulation source code, execution pipelines, nominal parameter manifests, co-simulation logs, and computational reproducibility metadata are openly available in the GitHub repository at: https://github.com/dafnhe-thor-barros/TCAS-DVFS.

Acknowledgments

The authors thank the Graduate Program in Control and Automation Engineering (ProPECAut) and the Federal Institute of Espírito Santo (IFES) for institutional and computational support. This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

References

BERGMAN, Theodore L., Adrienne S. LAVINE, Frank P. INCROPERA, and David P. DeWITT. Fundamentals of Heat and Mass Transfer. 7th ed., John Wiley & Sons, 2011.

BLACK, James R. "Electromigration—A Brief Survey and Some Recent Results." IEEE Transactions on Electron Devices, vol. 16, no. 4, 1969, pp. 338–347. DOI: 10.1109/T-ED.1969.16754.

BOYD, Stephen, and Lieven VANDENBERGHE. Convex Optimization. Cambridge University Press, 2004.

BOYD, Stephen, Neal PARIKH, Eric CHU, Borja PELEATO, and Jonathan ECKSTEIN. "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers." Foundations and Trends in Machine Learning, vol. 3, no. 1, 2011, pp. 1–122. DOI: 10.1561/2200000016.

BROOKS, David M., and Margaret MARTONOSI. "Dynamic Thermal Management for High-Performance Microprocessors." Proc. 7th International Symposium on High-Performance Computer Architecture (HPCA), 2001, pp. 171–182. DOI: 10.1109/HPCA.2001.903261.

CHEN, Haonan, Yuben QU, Tao WU, Jianchao ZHENG, Chao DONG, and Qihui WU. "Edge Intelligence for Space-Air-Ground Integrated Network: Challenges, Solutions, and Future Directions." Chinese Journal of Aeronautics, vol. 39, no. 8, 2026, art. 103969. DOI: 10.1016/j.cja.2025.103969.

CHEN, Shuyi, Zhengchang HUA, Nikos TZIRITAS, and Georgios THEODOROPOULOS. "Hot AI in Cold Space: Thermal-Crosstalk-Aware Scheduling for Sustainable Orbital AI Clusters." Proc. 3rd Workshop on Sustainable Computer Systems (HotCarbon), 2026, pp. 1–8. arXiv:2606.26150. DOI: 10.48550/arXiv.2606.26150.

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

DENBY, Bradley, and Brandon LUCIA. "Orbital Edge Computing: Nanosatellite Constellations as a New Class of Computer System." Proc. 25th International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS '20), ACM, 2020, pp. 939–954. DOI: 10.1145/3373376.3378473.

EUROPEAN COOPERATION FOR SPACE STANDARDIZATION (ECSS). Space Engineering: Thermal Control General Requirements. ECSS-E-ST-31C, Noordwijk, Netherlands, 2008.

GARRETT, Tyler, Seth ROFFE, and Alan D. GEORGE. "Soft-Error Characterization and Mitigation Strategies for Edge Tensor Processing Units in Space." IEEE Transactions on Aerospace and Electronic Systems, vol. 60, no. 4, 2024, pp. 5481–5498. DOI: 10.1109/TAES.2024.3393929.

GE, Yang, Parth MALANI, and Qinru QIU. "Distributed Task Migration for Thermal Management in Many-Core Systems." Proc. 47th ACM/IEEE Design Automation Conference (DAC), 2010, pp. 579–584. DOI: 10.1145/1837274.1837417.

GILMORE, David G., editor. Spacecraft Thermal Control Handbook: Fundamental Technologies. 2nd ed., The Aerospace Press / AIAA, 2002.

GIUFFRIDA, Gianluca, et al. "The Φ-Sat-1 Mission: The First On-Board Deep Neural Network Demonstrator for Satellite Earth Observation." IEEE Transactions on Geoscience and Remote Sensing, vol. 60, 2022, art. 5517414. DOI: 10.1109/TGRS.2021.3125567.

HOFFMANN, José Luis Conradi, and Antônio Augusto FRÖHLICH. "Online Machine Learning for Energy-Aware Multicore Real-Time Embedded Systems." IEEE Transactions on Computers, vol. 71, no. 2, 2022, pp. 493–505. DOI: 10.1109/TC.2021.3056070.

KANG, Minseon, and Moonju PARK. "Power Estimation and Energy Efficiency of AI Accelerators on Embedded Systems." Energies, vol. 18, no. 14, 2025, art. 3840. DOI: 10.3390/en18143840.

KING, Michael D., et al. "Spatial and Temporal Distribution of Clouds Observed by MODIS Onboard the Terra and Aqua Satellites." IEEE Transactions on Geoscience and Remote Sensing, vol. 51, no. 7, 2013, pp. 3826–3852. DOI: 10.1109/TGRS.2012.2227333.

LIN, Jai-Ming, et al. "Thermal-Aware Floorplanning and TSV-Planning for Mixed-Type Modules in a Fixed-Outline 3-D IC." IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 29, no. 9, 2021, pp. 1652–1664. DOI: 10.1109/TVLSI.2021.3100343.

MANSON, Samuel S. Thermal Stress and Low-Cycle Fatigue. McGraw-Hill, 1966.

MURALI, Srinivasan, et al. "Temperature-Aware Processor Frequency Assignment for MPSoCs Using Convex Optimization." Proc. 5th IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis (CODES+ISSS), 2007, pp. 111–116. DOI: 10.1145/1289816.1289845.

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION (NASA). Handbook for Models and Simulations: An Implementation Guide for NASA-STD-7009B. NASA-HDBK-7009B, Washington, DC, 2026.

PALLIPADI, Venkatesh, and Alexey STARIKOVSKIY. "The OnDemand Governor: Past, Present, and Future." Proc. Linux Symposium, vol. 2, 2006, pp. 215–230.

PEDERSON, Karl, et al. "Multifunctional Lightweight Radiators for Small-Satellite Thermal Control." Journal of Spacecraft and Rockets, 2026, art. A36693. DOI: 10.2514/1.A36693.

REUTHER, Albert, et al. "AI and ML Accelerator Survey and Trends." Proc. IEEE High Performance Extreme Computing Conference (HPEC), 2022, pp. 1–10. DOI: 10.1109/HPEC55821.2022.9926331.

SHERWOOD, R., et al. "The Autonomous Sciencecraft Experiment." Proc. 2003 IEEE Aerospace Conference, vol. 1, 2003, pp. 1-390–1-398. DOI: 10.1109/AERO.2003.1235068.

CC BY

REMUNOM
ISSN 2178-6925

Received: 14/10/2026 - **Accepted:** 12/11/2026
Vol: 13.23
DOI: 10.66104/v13w3915
Pages: 1-34

SIMON, Dan. Optimal State Estimation: Kalman, H-infinity, and Nonlinear Approaches. John Wiley & Sons, 2006.

SKADRON, Kevin, et al. "Temperature-Aware Microarchitecture: Modeling and Implementation." ACM Transactions on Architecture and Code Optimization (TACO), vol. 1, no. 1, 2004, pp. 94–125. DOI: 10.1145/980152.980157.

STELLATO, Bartolomeo, et al. "OSQP: An Operator Splitting Solver for Quadratic Programs." Mathematical Programming Computation, vol. 12, no. 4, 2020, pp. 637–672. DOI: 10.1007/s12532-020-00179-2.

WANG, Da-Wei, et al. "A Proposal of Fast Thermal Simulation Method for 2.5-D Advanced Packaging to Enable Efficient Thermal-Aware Placement Optimization." IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 45, no. 2, 2026, pp. 564–575. DOI: 10.1109/TCAD.2025.3585802.

YUAN, Zhehu, et al. "Heat: Satellite's Meat is GPU's Poison." arXiv preprint arXiv:2501.14757, 2025. DOI: 10.48550/arXiv.2501.14757.

ZHANG, Yufu, and Ankur SRIVASTAVA. "Accurate Temperature Estimation Using Noisy Thermal Sensors." Proc. 46th ACM/IEEE Design Automation Conference (DAC), 2009, pp. 472–477. DOI: 10.1145/1629911.1630036.

CC BY