

Settlement thermal engineering: heat rejection and recovery



Conceptual visualization — thermal control moves heat from sources to sinks and ultimately rejects it; radiator sizing depends strongly on temperature and emissivity.

Manage heat at settlement scale where habitats, crops, workshops, batteries and computing become thermally coupled.

Before starting — review thermodynamics, power and habitat interfaces because thermal rejection is a network problem spanning generation, transport, storage and final sinks. Every important symbol is defined at first use.

Mastery objectives

- draw the complete heat path from equipment and crew to transport loops, storage, recovery and final rejection
- separate installed, available and verified thermal capacity under N-1 and maintenance states
- calculate rejection, transport and transient margins without treating cold ambient conditions as automatic safety
- match recovered heat quality to useful sinks before deciding what still has to be rejected

1. Nearly all electrical power ends as heat

A settlement consumes electricity for lighting, computing, pumping, heating, cooling and manufacturing. Much of that energy eventually appears as heat inside equipment or occupied volumes. Thermal balance therefore tracks dissipated power, solar input, ground exchange and rejection capacity. A several-hundred-kilowatt electrical architecture is also a thermal architecture of comparable scale.

NASA Moon to Mars Architecture treats power and surface systems as coupled functions.

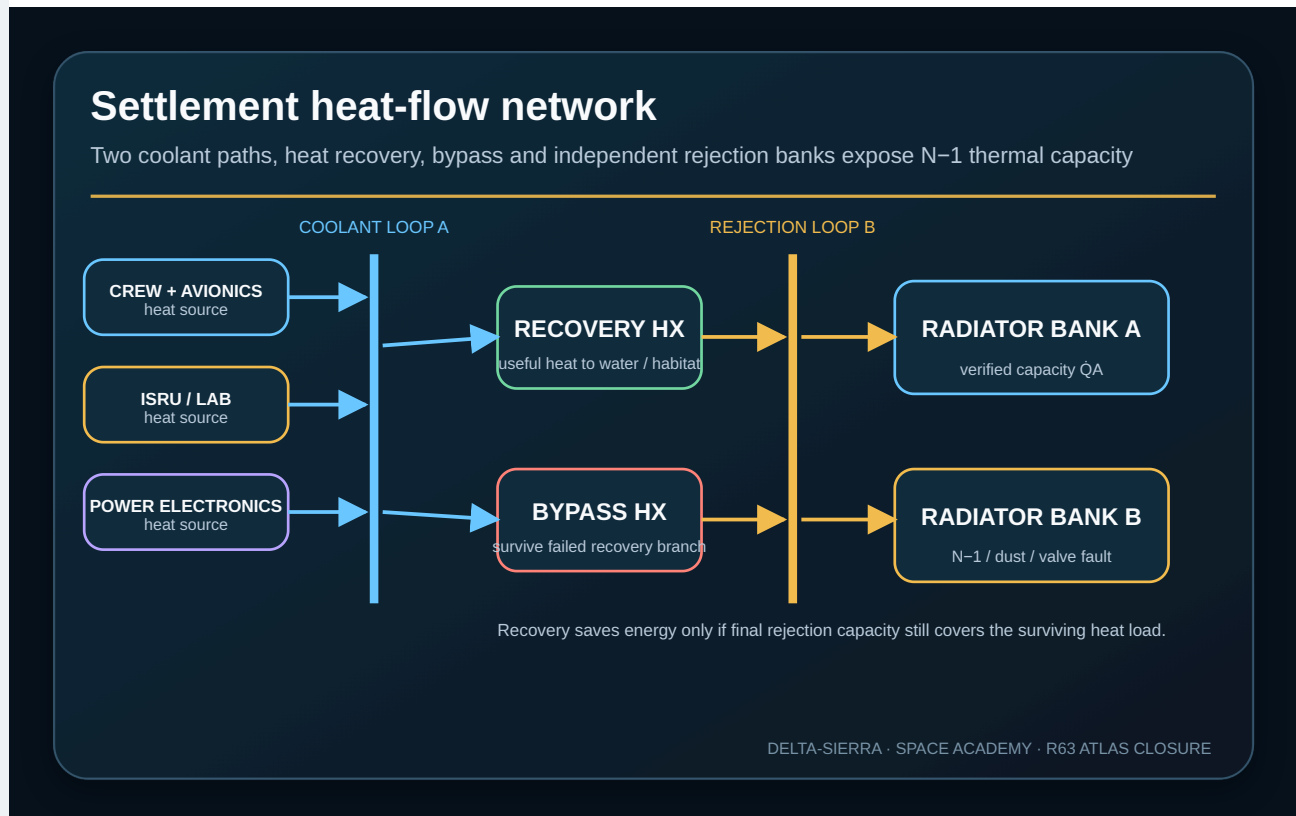
[NASA — Moon to Mars Architecture](#)

2. Temperature is not energy

Temperature describes state while heat describes energy transfer. A tiny component can be very hot yet contain little total energy, whereas a large water inventory with a small temperature rise can store many kilowatt-hours. Thermal sizing uses heat capacity, flow and temperature difference rather than degrees Celsius alone.

3. Fluid loops and heat exchangers

Active loops move heat from loads toward exchangers and radiators. Flow rate, specific heat and temperature rise determine transported power. Too little flow limits heat removal; excessive flow consumes pump power and can increase wear. Heat exchangers also need isolation and tolerance for fouling or leakage.



Thermal control succeeds only if every watt has a credible path from its source to a final sink. Pedagogical synthesis by Delta-Sierra from the primary sources cited in this course; schematic, not to scale.

4. Radiator area and operating temperature

Radiator performance depends strongly on absolute temperature, emissivity and view to a cold environment. Dust, shading and orientation reduce capacity. Raising loop temperature can reduce required area but may exceed equipment or material limits. Radiator sizing is therefore coupled to acceptable operating temperatures.

5. Recover heat before rejecting it

A greenhouse, dryer, water tank or industrial process can use heat that would otherwise be rejected. Recovery improves system efficiency only when temperature levels match. Thirty-degree waste heat cannot directly substitute for an 800 °C furnace. Thermal integration must consider the quality of heat, not only total kWh.

6. Thermal storage and time shifting

Water, phase-change material or structural mass can absorb heat for hours. Storage can smooth peaks and reduce instantaneous rejection size. It creates no energy sink, however; the stored heat still has to be used or rejected later. The architecture must close the full daily balance.

7. Thermal failures can become critical quickly

A computer or battery can remain electrically powered while exceeding temperature limits after coolant flow stops. Contingency design therefore estimates time-to-limit, loads to shed, passive circulation and backup-loop capacity. Trend alarms are more useful than waiting for a final high-temperature threshold.

8. Integrate habitat, crop and industrial heat

Crops may require heat at one time and produce excess at another. Workshops create peaks. Batteries need a narrow range. A mature settlement treats these as a thermal network with priorities and exchangers instead of independent air conditioners.

Deepening: coupling power and thermal contingency modes

Shedding an electrical load often reduces heat that must be rejected, but not always. Turning off a pump to save electricity can create rapid overheating. Emergency procedures therefore need simultaneous power and thermal budgets. Critical-load lists include equipment that removes heat from other loads, not only equipment whose function is directly visible to the crew.

Deepening: dust and radiator degradation

Martian dust can reduce radiator performance by changing emissivity or exposure. The system therefore includes inspection, performance trending and possibly cleaning. Degradation is detected by comparing rejected power with expected temperatures and conditions. Temporary compensation may include load shedding or a higher loop temperature, but maintenance must occur before thermal margin disappears.

Deepening: temperature hierarchy

Not every loop should operate at the same temperature. A cold loop can protect sensitive electronics while a warmer loop rejects industrial heat more effectively. Connecting all loads to one temperature level can oversize radiators and pumps. A multi-loop architecture exchanges heat between levels when useful and rejects the remainder at the warmest temperature compatible with equipment.

9. Worked example: coolant flow

A loop must carry 50 kW with water and allows a 5 K temperature rise. Using specific heat about 4.18 kJ/(kg·K), mass flow is $\dot{m} = 50 \div (4.18 \times 5) \approx 2.39$ kg/s. If the pump delivers only 1.5 kg/s, either the temperature rise must increase or less heat can be removed.

Calculated case study: recovered heat and heat still requiring rejection

TEACHING ASSUMPTION — Equipment dissipates 80 kW of heat. A recovery system can use 25% of that thermal power over a 10 h period.

Let P_d be dissipated thermal power in kW; η recovered fraction, dimensionless; P_r recovered power; P_j power still to reject; t duration in h; and E_j rejected thermal energy in kWh thermal.

$P_r = P_d \times \eta = 80 \times 0.25 = 20$ kW. $P_j = 80 - 20 = 60$ kW. Over 10 h: $E_j = P_j \times t = 60 \times 10 = 600$ kWh thermal.

Heat recovery reduces rejection duty but does not remove it. Real sizing depends on temperatures, heat exchangers and the radiative environment; the teaching goal is to close the power balance.

10. Exercise

A computing zone dissipates 35 kW and a workshop 60 kW for two hours. Combine a 70 kW radiator, thermal storage and load shifting. Calculate excess energy during the peak.

11. Reasoned solution

Combined load is 95 kW. With 70 kW rejection, the deficit is 25 kW for two hours, equal to 50 kWh thermal. Storage must exceed that value with margin, or workshop activity must be shifted.

12. Validation project

Build the thermal architecture for a habitat, crop area and workshop: loads, loops, exchangers, radiators, storage, heat recovery, failure cases and shedding plan.

A settlement is a heat-flow network as much as a power network

Almost every watt used inside a Mars settlement ultimately appears as heat. Computers, pumps, lighting, food processing, exercise equipment, life-support machinery and industrial tools convert electrical or chemical energy into thermal energy. Solar gains and heat moving through walls add another layer. The settlement must therefore decide where heat is useful, where it must be transported and where it must finally be rejected to the environment.

Thermal control is easy to underestimate because temperature changes can lag behind power changes. A habitat may appear stable while heat accumulates in equipment, coolant or structure. Conversely, a cold-soaked component can demand a large heating load even when the average habitat temperature is comfortable. Thermal engineering tracks heat rates, temperatures, storage capacity and transport paths separately.

Mars adds operational constraints. The thin atmosphere provides little convective cooling compared with Earth, so external heat rejection relies strongly on radiation and conduction paths. Dust can change optical properties. Long nights, seasonal changes and local geometry affect environmental temperatures. Thermal design therefore belongs in the same integrated model as power generation, life support, habitat layout and maintenance.

Ten thermal concepts to master

1. Internal electrical power becomes a thermal load

If a 2 kW computer rack runs inside a habitat, most of those 2 kW must eventually be removed as heat unless the energy leaves in another form. The same is true for motors and power electronics. A power budget that ignores the resulting thermal load is incomplete.

Internal sources should be mapped by location and duty cycle. A short high-power tool can create a local hotspot even when daily average energy is modest. Thermal design therefore needs both steady-state and transient cases.

2. Temperature is not the same thing as thermal energy

Temperature describes a thermodynamic state; heat is energy transferred because of temperature difference. A large water tank can absorb substantial heat with only a modest temperature rise because it has high heat capacity. A small electronic component can become dangerously hot with far less stored energy because its mass is small and its heat-transfer path is poor.

This distinction explains why thermal storage can smooth peaks. The goal is not to “store temperature” but to store energy in a material whose allowable temperature range and heat capacity are known.

3. Conduction controls many local bottlenecks

Heat moves through walls, structural members, mounting plates and thermal straps. A narrow or poorly conducting interface can dominate component temperature even when the external radiator is large. Thermal verification must therefore trace the entire path from heat source to final sink.

Insulation works by reducing heat transfer, which is beneficial for a warm habitat but can be harmful around equipment that must reject heat. The correct design creates low-conductance paths where heat should stay and high-conductance paths where heat should move.

4. Fluid loops move heat between zones

Pumped liquid loops can collect heat from equipment and deliver it to heat exchangers, storage tanks or radiators. Flow rate, fluid heat capacity and allowable temperature rise determine how much heat a loop can carry. Redundant pumps or cross-ties can prevent one pump failure from disabling an entire habitat.

Leaks, gas bubbles, freezing risk, contamination and pump wear become maintenance concerns. A loop should therefore be segmented and instrumented so the crew can isolate a fault and continue in a degraded mode.

5. Radiators trade area against temperature

Radiative heat rejection increases strongly with absolute temperature. A warmer radiator can reject more heat per square meter, but equipment and fluids may limit allowable temperatures. Surface emissivity, view to the sky, dust and nearby warm structures also affect performance.

Radiator placement should protect surfaces from vehicle traffic and landing plume contamination. Cleanability and inspection access matter because a settlement needs heat rejection for years, not for one mission demonstration.

6. Route useful heat to sinks before final rejection

Waste heat from electronics, industrial processes or greenhouse lighting can preheat water, warm an adjacent zone or support process temperature. Heat recovery reduces the energy that would otherwise be spent on heating while also reducing radiator load.

However, recovered heat must match a real demand at an appropriate temperature. Low-grade heat cannot replace every high-temperature process. A thermal hierarchy classifies sources and sinks by temperature so useful matches are made before the remaining heat is rejected.

7. Thermal storage decouples generation from rejection

A water reservoir or phase-change material can absorb heat during a short peak and release it later when radiator capacity is available. This is useful when industrial equipment runs intermittently or when the power system changes operating mode.

Storage is finite. Operators must know state of thermal charge, maximum allowable temperature and recovery time. Otherwise a buffer can mask an overload until it is nearly full, leaving little time to react.

8. Thermal faults can escalate quickly

A power shortage may allow non-critical loads to be turned off. A cooling-loop failure, by contrast, can overheat critical electronics or life-support equipment in minutes. Thermal fault detection should therefore monitor temperatures, flow, pressure and pump status with alarms tied to action thresholds.

Degraded modes can include reducing computing load, stopping industrial processes, moving crew between zones, opening cross-ties or switching to a backup radiator branch. These actions should be rehearsed and prioritized by consequence.

9. Habitat, greenhouse and industry should exchange heat deliberately

A greenhouse can be both a thermal load and a useful heat sink depending on lighting, crop cycle and temperature. Industrial areas may produce high-grade heat at times when habitats need heating. Integrated design seeks beneficial exchanges without coupling failures so tightly that one zone can disable another.

Interfaces should include isolation capability. Sharing heat is valuable during normal operation, but emergency operation may require a habitat to separate from a failed industrial loop.

10. Growth requires thermal capacity planning

Adding people and equipment increases heat generation. A settlement that has enough radiator capacity for thirty people may become thermally constrained before it becomes electrically constrained when it grows. Expansion plans should reserve piping, pump, exchanger and radiator capacity rather than treating thermal control as an afterthought.

Capacity planning should use peak simultaneous cases and realistic derating. Dust, partial radiator shading, a failed pump or seasonal environment can reduce available rejection below nameplate values.

Calculation laboratory

Internal heat generation

$$\dot{Q}_{\text{internal}} = \sum P_{\text{dissipated}}$$

The dot over Q means a rate of heat transfer, measured in watts. If three internal loads dissipate 2.0 kW, 1.5 kW and 0.8 kW, the total internal heat rate is 4.3 kW. This is the heat that thermal control must transport or temporarily store if the loads operate simultaneously.

Solar absorption

$$\dot{Q}_{\text{solar}} = \alpha \times A \times G$$

α is absorbed fraction, A is illuminated area and G is incident solar flux. The equation is a first-order estimate; orientation, shadows and surface properties matter. It is useful for understanding why dark external surfaces can create different thermal behavior from reflective surfaces.

Conduction through a simple layer

$$P_{\text{cond}} = k \times A \times \Delta T / L$$

k is thermal conductivity, A area, ΔT temperature difference and L layer thickness. If insulation is thicker, conductive heat transfer falls in this simplified one-dimensional model. Real structures include joints, penetrations and thermal bridges that must be assessed separately.

Heat carried by a liquid loop

$$\dot{Q}_{\text{fluid}} = \dot{m} \times c_p \times \Delta T$$

For water-like coolant, suppose mass flow is 0.10 kg/s, heat capacity is approximately 4.18 kJ/(kg·K), and coolant warms by 8 K. Heat transport is $0.10 \times 4.18 \times 8 = 3.344$ kJ/s, or about 3.34 kW. This relation lets engineers choose flow or temperature rise for a required transport load.

Radiative heat rejection

$$P_{\text{rad}} = \varepsilon \times \sigma \times A \times (T^4 - T_{\text{env}}^4)$$

ε is emissivity, σ the Stefan-Boltzmann constant, A radiator area and temperatures are absolute kelvin. The fourth-power relationship means radiator performance changes strongly with temperature. It also explains why thermal design must use kelvin rather than Celsius inside this equation.

Thermal storage

$$Q_{\text{store}} = m \times c_p \times \Delta T$$

A 500 kg water buffer allowed to warm by 10 K stores approximately $500 \times 4.18 \times 10 = 20,900$ kJ, or about 5.8 kWh of thermal energy. That may cover a short peak, but it is not a long-term heat sink. The radiator must eventually remove the stored energy.

Heat-recovery efficiency

$$\eta_{\text{recovery}} = Q_{\text{useful}} / Q_{\text{available}}$$

If 12 kWh of recoverable heat is available and 7.2 kWh is successfully delivered to useful sinks, recovery efficiency is 60%. The remaining heat still needs rejection or storage. A high percentage is useful only if the receiving process genuinely needs the recovered heat.

Thermal capacity margin

$$\text{Margin} = (P_{\text{capacity}} - P_{\text{peak}}) / P_{\text{peak}}$$

If available heat-rejection capacity is 60 kW and planned peak load is 48 kW, margin is $(60-48)/48 = 0.25$, or 25%. If dust or a failed branch derates capacity to 50 kW, effective margin falls to about 4.2%. This is why nominal margin must be tested against degraded states.

Worked settlement case: reuse heat before radiator sizing

A habitat and workshop together generate 42 kW of peak internal heat. During the same period, water processing and a greenhouse can accept 12 kW of useful recovered heat. The remaining 30 kW must be stored temporarily or rejected. If a thermal buffer can absorb 6 kWh, it can cover only twelve minutes at a 30 kW net load if no radiator is operating. This demonstrates that storage buys response time; it does not replace continuous rejection.

Suppose radiators can reject 36 kW in the expected condition. The system has 20% margin over the 30 kW net heat. If dust or geometry reduces radiator performance by 15%, available rejection falls to 30.6 kW, leaving almost no margin. Operators would then schedule workshop loads differently, improve radiator condition or activate another heat sink. The coupled power-and-thermal plan should define that degraded operating mode in advance.

Progressive exercises with solutions

- ▶ Exercise 1 - Fluid-loop heat transport
 - ▶ Exercise 2 - Thermal storage
 - ▶ Exercise 3 - Capacity margin
 - ▶ Exercise 4 - Degraded radiator
-

Interactive beginner glossary

- heat rate - thermal power flowing into or out of a system.
- conduction - heat flow through solid or contacting materials.
- thermal radiation - heat emitted as electromagnetic energy.
- radiator - external heat-rejection surface.
- coolant loop - pumped fluid network carrying heat.
- heat exchanger - device coupling two thermal circuits.
- thermal storage - temporary buffer for heat.
- derating - reduced allowable or available performance.

Design the thermal system as a network of heat paths

A beginner often imagines thermal control as a collection of heaters and radiators. An engineer instead starts with a map of heat sources, thermal masses, transport paths, controlled temperatures and final sinks. Every electrical device eventually becomes a heat source unless energy leaves the controlled volume in another form. Human metabolism, lighting, pumps, batteries, avionics, food preparation, laboratories, motors and chemical processes therefore participate in the thermal budget even when their primary purpose is not “thermal.”

The first design question is not “how large should the radiator be?” but “where is heat generated, at what temperature, how variable is it, and which equipment can tolerate which temperature range?” A warm electronics rack may be able to reject heat to a fluid loop at a temperature that is useful for preheating another process. A greenhouse may have a different temperature band from a battery room. A medical bay may demand narrow environmental stability. Treating all heat as one interchangeable stream can hide these constraints.

Steady state and transient state are different problems

In steady state, average heat entering a controlled node equals average heat leaving it. During a transient, the difference changes the stored thermal energy of the node. A habitat can therefore survive a short mismatch by allowing a component, fluid tank or structure to warm or cool within its allowable band. This thermal inertia is useful, but it is not free capacity: once the temperature limit is reached, the imbalance must be removed.

Thermal storage during a transient

$$Q_{\text{stored}} = m \times c_p \times \Delta T$$

Read aloud. Stored sensible heat equals mass times specific heat capacity times temperature change.

For a teaching example, take 500 kg of a fluid-like thermal mass with an illustrative specific heat capacity of $4.0 \text{ kJ} \cdot \text{kg}^{-1} \cdot \text{K}^{-1}$ and allow a 5 K rise. Stored energy is $500 \times 4.0 \times 5 = 10,000 \text{ kJ}$, or 10 MJ. If an excess heat load is 10 kW, which is 10 kJ/s, the idealized buffer would last $10,000/10$

= 1,000 s, about 16.7 minutes. Real usable time would depend on mixing, local hot spots, pump operation and allowable component temperatures.

Temperature level determines whether “waste heat” is actually useful

Heat recovery is constrained by temperature. A warm loop can transfer energy to something colder, but not every low-temperature heat source can efficiently satisfy a higher-temperature process. The engineering map should therefore record not only kilowatts but also source and sink temperatures. Two streams with the same heat rate may have very different usefulness.

This matters for settlement architecture because heat reuse can reduce simultaneous heating demand while increasing plumbing, controls and failure coupling. A heat exchanger can connect two previously independent systems. If contamination, leakage or control failure in one loop can affect the other, the energy saving must be traded against the new dependency. The safest architecture is not automatically the one with the highest recovery percentage.

Dust and geometry are operational variables

Radiator performance depends on view geometry and surface condition. A radiator that works well in a clean analytical model can perform differently when partially shadowed, contaminated or covered by dust. A Mars settlement therefore needs inspection criteria, cleaning strategy where appropriate, temperature sensors distributed enough to reveal non-uniform performance, and a degraded operating mode that sheds nonessential heat sources before protected components exceed limits.

Thermal control should also be reviewed against maintenance access. A perfectly sized radiator that cannot be isolated without shutting down the whole habitat is operationally weak.

Likewise, a coolant loop that requires a rare pump seal or specialized tool can create a logistics dependency larger than its mass suggests.

Failure drill: rejection capacity falls while internal load rises

Consider a teaching scenario in which one radiator branch becomes unavailable just as a workshop begins a power-intensive repair. The correct response is not to calculate only the lost radiator area. The team identifies which zones are warming, how quickly temperatures are

moving, what thermal mass is available, which loads can be paused, whether another loop can accept heat, and what restart conditions will prove stability.

The operations display should show a *time to thermal limit* for the most constrained component rather than a single green/red thermal status. If a battery enclosure has 40 minutes to its limit while a greenhouse has several hours, the response order is clear even before the fault is fully diagnosed.

Engineering exercise: build a heat-path table

Create a table for ten settlement loads. For each, record nominal power, fraction ultimately appearing as heat inside the controlled volume, normal temperature band, thermal transport path, final rejection path, sensors, isolation method, backup path and time to unacceptable temperature after loss of cooling. Then remove one common pump or power bus from the diagram and identify every load whose thermal control fails because of that single dependency.

The exercise is passed only when the student can explain why an apparently simple change in electrical load can alter radiator demand, why thermal storage buys time but not indefinite survival, and why heat recovery can create both efficiency and coupling.

Thermal calculation laboratory: radiators are governed by the fourth power of temperature

Heat can be transported by conduction and fluid loops inside a system, but in space the final rejection path is primarily radiation. That makes radiator temperature a powerful design variable.

Net radiative heat rejection

$$\dot{Q} = \epsilon \sigma A (T^4 - T_{\text{sink}}^4)$$

Read aloud. Heat-rejection rate equals emissivity times the Stefan–Boltzmann constant times radiator area times the difference between the fourth powers of radiator and effective sink temperature.

- $\dot{Q}$: heat-rejection rate, watts.
- ϵ : emissivity, dimensionless.
- σ : Stefan–Boltzmann constant, approximately $5.67 \times 10^{-8} \text{ W} \cdot \text{m}^{-2} \cdot \text{K}^{-4}$.
- A : effective radiating area, m^2 .
- T and T_{sink} : absolute temperatures in kelvin.

Teaching scenario. With $\epsilon = 0.9$, $A = 20 \text{ m}^2$, $T = 300 \text{ K}$ and an effective sink of 210 K , the idealized calculation gives about 6.28 kW net rejection. The fourth power means raising radiator temperature can strongly increase rejection, but component temperature limits and working-fluid choices constrain that option.

Limit. A real Mars surface radiator also sees the sky, Sun, ground, dust and changing geometry; this simplified expression is a first sizing lesson, not a final design.

Quantify recovered heat before sizing final rejection

$$\dot{Q}_{\text{useful}} = \eta_{\text{recovery}} \times \dot{Q}_{\text{waste}}$$

Apply the heat-recovery principle from the core lesson numerically: subtract only heat that can actually be transferred at a useful temperature and with an available exchanger. The remainder still needs a credible path to the radiator or another ultimate sink.

► Exercise — recovery rate

HEAT SOURCES

crew · electronics · process

COOLANT LOOP

collect · transport · isolate

RADIATOR

area · temperature · emissivity

SPACE / MARS

SKY

radiative sink · dust constraints

Engineering studio: size a thermal network from heat sources to the Martian sky

A thermal system is easiest to understand when the learner stops thinking of “temperature control” as one box and instead follows every watt of heat. Electrical equipment, people, lights, pumps, computers, food preparation and industrial processes release heat into different zones. Some of that heat can be moved to a useful sink, such as a greenhouse or a water-preheating loop. The remainder must eventually cross a heat exchanger, a coolant loop and a radiator before it can leave the settlement. On Mars there is no outside air dense enough to carry away large continuous loads by ordinary convection, so the final rejection step is dominated by radiation.

Begin by drawing a heat-path table with five columns: source, power in watts, useful temperature level, transport path and final sink. The purpose is not administrative bookkeeping. It reveals impossible designs. A habitat can have plenty of electrical generation and still overheat if its radiator loop is too small, its pump is unavailable, a valve isolates the wrong branch or a dust-loaded surface rejects less heat than assumed.

Node balance: conservation of energy before control theory

Take one thermal node, such as an avionics bay or inhabited room. Over a chosen time interval, heat entering the node plus heat generated inside it must equal heat leaving the node plus any heat stored by increasing the node temperature. That statement is conservation of energy. It is the foundation of every more sophisticated thermal model.

Transient node balance

$$m \times c_p \times dT/dt = \dot{Q}_{in} + \dot{Q}_{gen} - \dot{Q}_{out}$$

Question. How fast will a room or equipment mass warm when heat arrives faster than the cooling system removes it?

Read aloud. Mass multiplied by specific heat capacity multiplied by the rate of temperature change equals incoming heat plus internally generated heat minus outgoing heat.

Symbols and units. m is effective thermal mass in kilograms; c_p is specific heat capacity in joules per kilogram-kelvin; dT/dt is temperature change in kelvins per second; every $\dot{Q}$ term is thermal power in watts, where one watt is one joule per second.

Teaching calculation. Suppose an effective 2,000 kg equipment-and-structure mass has $c_p = 800 \text{ J}/(\text{kg}\cdot\text{K})$. A cooling fault leaves a net excess heat of 12 kW. The heat capacity is $2,000 \times 800 = 1,600,000 \text{ J/K}$. The warming rate is $12,000 / 1,600,000 = 0.0075 \text{ K/s}$, about 0.45 K/min. A 10 K allowable rise would therefore be reached in roughly 22 minutes if the simplified assumptions remained valid.

Independent check. Twelve kilowatts supplies 720 kJ each minute. Dividing 720 kJ by 1,600 kJ/K gives 0.45 K/min, the same result.

Limit. Real rooms exchange heat with walls, adjacent zones, ventilation, equipment housings and phase-change materials. The simplified node estimate is useful for urgency, not a replacement for a validated transient model.

Pump power is part of the thermal budget

Moving heat also consumes power, and that power ultimately becomes additional heat. A liquid loop must overcome pressure losses in pipes, bends, valves, filters and heat exchangers. If a designer sizes the electrical system and radiator system separately, the loop can hide this coupling.

Hydraulic pump power

$$P_{\text{pump}} \approx \Delta p \times \dot{V} / \eta$$

Question. How much electrical power is required to drive a coolant flow against a pressure drop?

Symbols. Δp is pressure drop in pascals; $\dot{V}$ is volumetric flow in cubic metres per second; η is pump efficiency expressed as a fraction; P is watts.

Example. For $\Delta p = 120,000 \text{ Pa}$, $\dot{V} = 0.003 \text{ m}^3/\text{s}$ and $\eta = 0.65$, hydraulic input is $120,000 \times 0.003 = 360 \text{ W}$. Electrical input is $360/0.65 \approx 554 \text{ W}$. That extra electrical power becomes heat somewhere in the system, so it cannot be ignored in a closed thermal budget.

Sanity check. Because efficiency is below one, electrical input must be larger than useful hydraulic power. A result below 360 W would signal an error.

Design failure states, not only nominal state

For each radiator branch, ask what happens if one pump stops, one loop is isolated, dust reduces optical performance, a coolant leak empties one circuit, a valve freezes, or an electrical bus is shed. A robust architecture may use sectionalized loops, cross-ties and protected power for minimum heat rejection. The learner should state which equipment can be shut down first, which rooms can tolerate a wider temperature band and which biological or medical loads cannot.

Integrated thermal drill

A settlement normally rejects 180 kW . One of three equal radiator branches is isolated and internal emergency activity adds 15 kW . First estimate the missing rejection capacity if branches share the load equally. Then identify at least four operational responses that reduce net heat production or move heat into temporary storage. Finally state the measurement that would tell the crew whether the situation is stabilising.

► Reasoned solution

First-Man thermal desk: heat must always have somewhere to go

Thermal control is easy to underestimate because heat is invisible. A habitat can have adequate electrical power, oxygen and water yet still become uninhabitable if metabolic heat, electronics, lighting, chemical processing and solar gains cannot be transported and rejected. The central habit for a thermal engineer is therefore to draw the entire heat path: where heat is created, where it can be stored temporarily, how it is transported, and where it finally leaves the system.

On Mars, “outside is cold” does not mean cooling is free. A low-density atmosphere offers limited convective heat transfer compared with Earth, dust can alter radiator performance, and many components must stay above minimum temperatures at the same time that other components need aggressive cooling. Heating and cooling are therefore coupled. Waste heat can sometimes be recovered usefully, but recovery is valuable only if the receiving load exists when the heat is available.

Turn a transient into a clock

$$\Delta t = C_{th} \Delta T / \dot{Q}_{net}$$

Starting question

If active heat rejection is lost, how long can a thermal mass absorb the imbalance before a temperature limit is reached?

Read aloud

Read: “time equals thermal capacitance times allowed temperature rise divided by net heat input.”

Symbols, pronunciation and meaning

C_{th} is effective thermal capacitance in kJ/K; ΔT is the allowed rise in kelvins; $\dot{Q}_{net}$ is net heat accumulation rate in kW, where 1 kW = 1 kJ/s; Δt is seconds.

Units

$(\text{kJ/K} \times \text{K}) \div (\text{kJ/s}) = \text{s}$. The kelvin cancels and the energy units cancel.

Origin and status of values

Thermal capacitance comes from masses and specific heats participating on the relevant timescale. Allowed ΔT comes

from equipment or crew limits. Net heat is generated heat minus whatever passive or residual rejection still works.

Why this operation

The numerator is how much thermal energy can be stored before the limit. Dividing that stored-energy allowance by the rate at which energy accumulates gives time.

Substitution and calculation

Teaching case: $C_{th}=18,000$ kJ/K, allowed rise 8 K, net accumulation 24 kW. Stored allowance = 144,000 kJ. Since 24 kW = 24 kJ/s, time = $144,000/24 = 6,000$ s = 100 min.

Calculator entry

Enter $18000 \times 8 \div 24$, then divide the result by 60 to convert seconds to minutes.

Mental estimate

$18,000 \times 8$ is about 144,000; dividing by roughly 25 gives slightly under 6,000 s, or around 100 min.

Independent check

Twenty-four kilojoules each second for 6,000 s equals 144,000 kJ, exactly the assumed thermal storage allowance.

Physical or operational interpretation

A 100-minute clock is not "100 minutes to start troubleshooting." Detection, diagnosis, crew access, valve reconfiguration and restart all consume part of it. An operational limit should therefore be shorter.

Plain-English translation

Thermal mass buys time; it does not remove heat. Once the storage margin is consumed, temperature keeps rising unless heat rejection is restored or heat production is reduced.

Variation / sensitivity

Halving the net heat load doubles the time. Cutting nonessential electrical loads can therefore be a powerful thermal emergency action.

Limit / assumption

This is a lumped-capacitance teaching model. Real systems have gradients, changing properties, multiple nodes, control actions and heat losses that vary with temperature.

What this does not prove

The calculation does not show that every component remains within limits; a local electronic box can overheat long

before the average habitat node reaches the selected temperature.

Boundary case to test

If $\dot{Q}_{\text{net}}$ approaches zero, the simplified time tends toward infinity because no net heat is accumulating. That mathematical limit signals a steady balance, not an infinitely robust system: pumps, sensors and local hotspots can still fail.

Radiator sizing is a fourth-power problem

Radiation grows strongly with absolute temperature. A radiator operated warmer can reject much more heat per square metre, but the coolant loop and connected equipment must tolerate that higher temperature. Conversely, a colder loop protects temperature-sensitive equipment but demands more radiator area. This trade is why “make the radiator bigger” and “run it hotter” are not equivalent design moves.

For a grey surface, a useful first model is $\dot{Q} = \epsilon \sigma A (T^4 - T_{\text{sink}}^4)$. The result depends on emissivity ϵ , Stefan–Boltzmann constant σ , area A and absolute temperatures in kelvins. The surrounding radiative sink is not simply the Martian air temperature; view factors to the sky, ground and nearby structures matter. A radiator facing warm sunlit terrain can behave differently from one viewing cold sky.

Design the degraded architecture before choosing the nominal setpoint

Ask what remains if one pump stops, one radiator panel is isolated, a dust event reduces rejection, a valve fails in place or a loop leaks. The degraded mode may shed greenhouse lighting, pause an industrial reactor, lower computing loads and prioritise crew cabin, medical storage and power electronics. Those priorities should be documented before the failure, because thermal time constants can turn a slow-looking event into a time-critical one.

Detect

temperature, flow,
pressure, pump current

Reduce heat

shed discretionary
electrical and process
loads

Reconfigure

alternate pump, bypass
or radiator branch

Verify

stable trend and
remaining margin

Worked design review

A settlement normally rejects 72 kW with three equal radiator branches. One branch is isolated after a leak. If the remaining two branches can each safely reject 30 kW at the current coolant temperature, available rejection is 60 kW, leaving a 12 kW deficit. The first response is not necessarily to raise the coolant temperature immediately. Operators can pause a 9 kW process and shed 4 kW of discretionary computing, turning the deficit into 1 kW of spare capacity while the thermal team evaluates whether a higher-temperature mode is acceptable.

The lesson is architectural: load shedding, thermal storage, branch isolation and repair access are part of thermal control. Radiator area alone does not create resilience.

Thermal redundancy is meaningful only when cross-ties, pumps, controls and radiator sections are actually available in the degraded configuration. Pedagogical synthesis by Delta-Sierra from the primary sources cited in this course; schematic, not to scale.

Thermal commissioning dossier: prove the heat path before population growth

A settlement should not treat thermal control as a component check. Commissioning must prove the complete path from source to sink under nominal and degraded conditions. A pump can run, a radiator can radiate and a control valve can move while the integrated system still fails because the heat is trapped in the wrong node or because the surviving route cannot carry the peak load after isolation.

Commission the network in layers

Begin with instrumentation: verify temperature sensors, flow meters, pressure measurements and valve positions against known states. Then test local loops with controlled heat loads. Next, demonstrate transfer between loops and final rejection. Only after these pieces are known should the team run combined habitat, greenhouse and industrial loads. The objective is not to create one impressive “full power” test but to build evidence that allows a later fault to be localised.

Commissioning records should preserve the configuration used for each result. If a later software update changes pump control, a radiator panel is added or coolant composition changes, the old evidence may no longer bound the new configuration. This is configuration control applied to thermal safety.

Verified heat-rejection margin

$$M_Q = (\dot{Q}_{\text{reject,verified}} - \dot{Q}_{\text{peak,protected}}) / \dot{Q}_{\text{peak,protected}}$$

1 – Concrete question

How much demonstrated heat-rejection capacity remains above the protected peak thermal load in the operating state being reviewed?

2 – Intuition

Compare what the tested radiator/loop architecture can reject with the heat that must be removed to keep protected equipment and inhabited zones within limits.

3 – Quantities

Use verified rejection capacity under the relevant environment and the protected peak load after any planned thermal load shedding.

4 – Formula

Margin equals verified rejection minus protected peak heat, divided by protected peak heat.

5 – Read aloud

“M Q equals Q-dot reject verified minus Q-dot peak protected, divided by Q-dot peak protected.”

6 – Symbols

$\dot{Q}$ means thermal power; M_Q is a dimensionless margin.

7 – Pronunciation

$\dot{Q}$ is read “Q dot.”

8 – Units

Use watts or kilowatts for every $\dot{Q}$ term.

9 – Convention

Define whether the verified capacity is nominal, N-1, dust-degraded or seasonal. Never mix states in one ratio.

10 – Why this relationship

The numerator is spare heat-rejection capacity; dividing by demand makes that spare comparable to the thermal load.

11 – Assumptions

The capacity is assumed sustainable for the required duration and compatible with pump, radiator and coolant limits.

12 – Unit check

$$(kW - kW) / kW = 1.$$

13 – Numerical case

Verified sustainable heat rejection: $\dot{Q}_{\text{reject,verified}} = 440 \text{ kW}$.

Protected peak heat load: $\dot{Q}_{\text{peak,protected}} = 360 \text{ kW}$.

Thermal surplus = $440 - 360 = 80 \text{ kW}$.

$$M_Q = 80 / 360 = 0.2222\dots$$

$$M_Q \approx 22.2\%.$$

14 – Operations

Subtract to find 60 kW spare, then divide by 250 kW.

15 – Algebra check

$$250 \times 1.24 = 310 \text{ kW}.$$

16 – Mental estimate

Sixty is just under one quarter of 250, so 24% is consistent.

17 – Interpretation

The network has 24% demonstrated headroom in this state; that headroom can be consumed by growth, warmer radiator conditions, dust or equipment degradation.

18 – What it does not prove

It does not prove every local node remains cool. A bottleneck heat exchanger can overheat one zone while total settlement rejection still looks adequate.

19 – Sensitivity

If dust and view-factor changes reduce verified rejection by 12%, capacity becomes about 273 kW and margin falls to about 9%.

20 – Practice

Guided exercise. Calculate verified thermal margin for 440 kW rejection capacity and 360 kW protected peak heat.

Detailed guided correction.

1. Surplus rejection capacity = $440 - 360 = 80$ kW.
2. $M_Q = 80 \div 360 = 0.2222\dots$
3. Thermal margin $\approx 22.2\%$.
4. The margin is valid only for the tested boundary and environment; it does not automatically cover a failed pump, dust-degraded radiator or hotter seasonal case.

Autonomous exercise. A three-loop settlement has verified loop capacities of 170, 150 and 140 kW against protected loads of 130, 110 and 100 kW. If loop 2 fails, cross-ties can transfer at most 45 kW to loop 1 and 35 kW to loop 3. Determine whether the 110 kW lost load can be fully protected.

► Autonomous correction – open after attempting the exercise

21 – Mission decision

Block new permanent heat-producing loads when demonstrated contingency margin falls below the programme criterion, even if electrical generation remains abundant.

Thermal load shedding must be designed

Electrical load shedding can reduce heat, but the mapping is not one-to-one. Turning off a pump may save electrical power while making local temperatures worse. Stopping a chemical process may require a controlled cooldown. A greenhouse can tolerate some lighting reduction yet still require air circulation and root-zone temperature control. The thermal emergency table therefore needs its own priorities, linked to but not copied from the electrical priority table.

Seasonal and transient envelopes

A system should be reviewed across environmental states, not one “Mars average.” Solar input, dust deposition, sky view, operating temperature and internal activity all alter the heat balance. Transients deserve a separate check because thermal mass can hide a shortage for minutes or

hours. A system that survives a short test because walls absorb heat may still be incapable of steady operation.

Qualification drill

Create a commissioning sequence for a settlement with habitat, greenhouse and workshop loops feeding two radiator fields. Demonstrate normal operation, one pump loss, one radiator-field isolation and a case in which workshop heat rises while the greenhouse asks for recovered heat. For every test, define the measured evidence that would permit a PASS and the configuration details that must be recorded so the result remains meaningful six months later.

Source context. NASA thermal-management material and Moon-to-Mars architecture provide the primary engineering context. Numerical capacities above are Delta-Sierra teaching assumptions. [NASA JSC — Thermal Management Subsystems](#).

Failure review: when “outside is cold” becomes a dangerous shortcut

Mars can be cold while a radiator still struggles. The final heat-rejection rate depends on radiator temperature, emissivity, view to cold surroundings and absorbed environmental heat. A radiator that is shadowed, dusted, poorly oriented or partially blocked by nearby structures may reject less heat than a simple ambient-temperature intuition suggests. Thermal design therefore belongs to geometry and surface condition as much as to weather.

Review three states separately: steady nominal operation, a short transient after a fault and the eventual degraded steady state. Thermal mass can make the first minutes look acceptable even when the degraded steady state is impossible. A useful commissioning test records not only temperatures but their rates of change. If a protected node keeps warming after the initial transient, the system has not reached a safe equilibrium.

Maintenance planning: preserve heat rejection before the margin disappears

Radiator panels, pumps, valves, filters and heat exchangers degrade in different ways. A preventive-maintenance plan should connect each degradation indicator to lost capacity. Rising loop pressure drop, reduced flow at the same pump command, growing approach temperature across a heat exchanger or a persistent rise in radiator outlet temperature can all signal changing performance. Trend data let the crew intervene before the system crosses a hard limit.

Spare strategy should include not only replaceable pumps but seals, compatible coolant, sensors, valve actuators, electrical drives and tools needed to isolate a branch. If a repair requires draining a loop, the inventory of replacement fluid and the ability to capture contaminated coolant become part of thermal resilience.

Design review question set

Before approving a new industrial load, ask: where does every added watt end up; which loop receives it; which heat exchanger transfers it; which radiator rejects it; what happens if one pump or radiator segment is unavailable; what transient temperature rise occurs before shedding; and what operator action is required? A project that can answer only the electrical-power side of those questions is not ready to connect to the settlement.

R59 thermal-failure board: follow heat from source to sink under degraded operation

Thermal control is easy to underestimate because heat is invisible and Mars is cold. Equipment, people, lighting, batteries, electronics and industrial processes still generate heat that must cross interfaces before it can be rejected or stored. The correct mental model is a path: source → local collection → transport loop → exchanger → final rejection surface or temporary thermal storage. A failure at any stage can limit the entire chain.

Separate capacity from transport

A radiator can have theoretical rejection capacity while a failed pump prevents heat from reaching it. Likewise, a healthy loop can move heat to a radiator whose surface is degraded or whose view factor is poor. Commissioning should therefore test temperatures, flows and pressure drops at the integrated operating point. A single “radiator capacity” number is insufficient evidence.

Thermal inertia buys time, not permission

Structures, coolant and stored material absorb heat, so temperatures may rise slowly after a rejection failure. This can create a valuable response interval, but it can also hide the loss until margin is small. Estimate hold-up time with conservative heat capacity and allowable

temperature rise, then verify with sensors and trend rate. The crew should know which loads must be shed before the protected temperature limit is approached.

N-1 claims require cross-tie realism

Parallel loops are not automatically mutually supporting. Cross-tie valves, pipe diameters, pumps, exchanger areas and control software constrain how much load can move after a failure. The autonomous exercise in the 21-step card deliberately exposes this: nominal spare capacity can exist while cross-tie limits leave protected heat unserved. N-1 capability should be demonstrated at representative load, not inferred from nameplate sums.

Maintenance can consume thermal margin

Dust, coating damage, degraded pumps, fouled exchangers, sensor drift and gas accumulation can erode performance gradually. Trend the relationship between load and temperature rather than waiting for a high-temperature alarm. A growing approach temperature across an exchanger or a pump operating farther from its expected point can become an early maintenance trigger.

Failure-injection drill

Run a table-top exercise in which one loop trips during greenhouse lighting and battery charging. Identify which heat sources can be deferred, which cannot, how cross-ties are configured, how long thermal inertia is credited, and which temperature or pressure trend forces escalation. The answer should include both the engineering calculation and the operational sequence: who sheds the load, who verifies valve position, who watches the trend and what constitutes recovery.

Primary-source bridge. NASA Johnson Space Center describes spacecraft thermal-management subsystems, while Moon to Mars architecture work treats infrastructure as an integrated system. The Mars-settlement cases here are educational scenarios. [NASA JSC — Thermal Management Subsystems](#).

R60 thermal operations board: the colony needs a heat-dispatch system, not just larger radiators

A mature settlement should operate its thermal network with the same discipline used for electrical dispatch. The question is not simply whether the installed radiator area can reject the average heat load. Operators need to know which loads are producing heat now, which loops are carrying it, which sinks remain available, which exchangers are fouled or isolated, how quickly temperatures are moving, and which actions reduce heat generation without accidentally disabling the equipment that removes heat. This creates a live thermal operating picture rather than a static sizing worksheet.

Separate installed, available and verified rejection capacity

Installed capacity is the nameplate capability of the complete architecture. Available capacity removes equipment that is intentionally offline, isolated or under maintenance. Verified capacity is narrower again: it is the heat-rejection capability that has been demonstrated under the present dust, orientation, loop-temperature and pump conditions. A settlement that mixes those three quantities can believe it has a comfortable margin while its actually demonstrated capacity is much smaller. The operations board should display all three and make the difference visible to the crew.

Primary-source bridge. NASA's Johnson Space Center thermal-control material describes heat acquisition, transport and rejection as linked spacecraft functions; the settlement-scale dispatch model here extends that systems logic to a larger Mars surface network. [NASA — JSC Thermal Management Subsystems](#).

N-1 must be stated for the thermal network itself

An N-1 claim is meaningful only if the failed item is named. Losing one pump, one radiator panel, one coolant branch, one control cabinet or one district heat exchanger can produce very different consequences. For each credible single failure, the team should identify the surviving heat path, the loads that must be shed, the time before a protected temperature limit is reached, and the maintenance action needed to restore redundancy. A design that survives loss of one radiator panel but fails after the loss of a common pump controller is not genuinely N-1 at the system level.

Primary source at use. NASA JSC thermal-management material treats heat acquisition, transport and rejection as an integrated function. That is the evidence bridge for testing loss of a pump, branch or rejection element at the actual network level. [NASA JSC Thermal Management Subsystems](#).

Thermal storage buys time, not immunity

Water tanks, phase-change material and massive structures can absorb short peaks, but stored heat becomes a deferred obligation. Operators therefore need two clocks: how long storage can prevent a temperature limit from being crossed, and how long the system will later need to discharge that stored heat while normal loads continue. A workshop peak shifted into storage can still create a night-time crisis if the radiator network is already occupied rejecting habitat and greenhouse loads.

Dust degradation needs trend detection before emergency response

Radiator degradation should be discovered by performance trending rather than by the first overtemperature alarm. For a repeatable operating point, compare loop inlet and outlet temperatures, flow, pump power, surface temperature and rejected heat against a reference envelope. A gradual divergence can indicate dust accumulation, loss of emissivity, flow restriction or sensor drift. The maintenance trigger should be reached while there is still enough thermal margin to clean, isolate or reconfigure without entering an emergency load-shed state.

Match recovered-heat temperature to the end use

This section applies the earlier recovery principle to temperature quality. Warm coolant that can preheat water may be useless for a high-temperature industrial process; a heat ledger therefore records temperature level, exchanger availability and destination as well as kilowatts.

Commission the network by forcing realistic degraded states

Before relying on the network for crew growth, test it with one pump unavailable, one radiator branch isolated, one sensor channel declared unreliable and one high-load industrial event scheduled. The point is not to create drama but to verify that the monitoring, automatic protection, manual procedures and staffing can keep protected loads inside their temperature limits. Record the actual time to detect the fault, identify the correct branch, execute the reconfiguration and stabilise the loop. These measured times belong in the contingency model.

Scenario exercise — heat rejection is sufficient on paper but not operationally

A district has 500 kW of installed rejection, 455 kW currently available and 430 kW verified under present dust conditions. Protected heat load is 390 kW. A 60 kW industrial process is scheduled for two hours. The correct decision is not “500 exceeds 450, proceed.” Operators should recognise that the verified total would become 450 kW against 430 kW of demonstrated rejection, producing a 20 kW deficit before any uncertainty or additional failure. The process should be delayed, reduced, coupled to qualified storage or run only after another load is shifted. This is the difference between installed capacity and operationally defensible capacity.

R60 qualification drill: prove the settlement can cross a thermal upset without spending every reserve

For qualification, choose one representative hot-day operating state and freeze the assumptions: habitat occupancy, greenhouse lighting, computing load, industrial schedule, radiator condition, coolant inventory and dust state. Then inject one credible loss, such as a pump trip or radiator branch isolation. The team should predict the first protected temperature to approach its limit, identify the load-shedding order, state the thermal storage available, and define the measurement that proves stabilisation. The exercise is failed if the team relies on “nominal radiator capacity” after the injected loss without recomputing the surviving path.

A second phase should begin after apparent stabilisation. Operators restore one discretionary load and observe whether the network can return to a sustainable steady state rather than merely postponing the temperature rise. This catches a subtle failure mode: a system may look recovered for twenty minutes because thermal mass absorbs the imbalance, while temperatures are still drifting toward a limit. Qualification therefore requires a trend criterion, not only a snapshot below the alarm threshold.

The maintenance team should then inspect the work required to restore redundancy. If recovering one radiator branch demands an EVA, a rare seal, a specific torque tool or the only thermal specialist, those dependencies belong in the readiness record. The thermal architecture is not fully resilient until the restoration path is supportable by the real crew, inventory and access constraints.