HyperNext
Data Centers

RESEARCH PAPER

HN-RP-006

Liquid Cooling at 600 kW per Rack

Direct-to-chip engineering for the rack-scale GPU era

Why air stops working. Glycol coolant chemistry and flow paths. The failure modes that matter.

This is what is Next.

Series	HyperNext Research
Paper	HN-RP-006
Issued	29 April 2026
Version	1.0
Classification	Public release
Citation	HyperNext Research, HN-RP-006

Liquid Cooling at 600 kW per Rack

This paper is part of the HyperNext Research series. Methodology, assumptions, and source data are stated openly so other operators can reproduce the analysis on their own facilities. Citation as "HyperNext Research, HN-RP-006" is welcome.

Contents

1. The crossover from air to liquid

ABSTRACT

Air cooling stops working as a primary thermal medium somewhere around 50 kilowatts per rack. From that crossover up to the 600 kilowatts that the Vera Rubin Ultra NVL576 will draw, every kilowatt has to be moved by liquid. This paper covers the mechanics of doing that at production scale. We work through the cold plate design, the rack-level coolant distribution, the row-level coolant distribution unit (CDU), the secondary chilled-water loop, and the heat exchanger that hands the heat off to the building cooling plant. We use propylene glycol coolant in the primary loop, not water. We explain why. We cover the failure modes specific to high-density direct-to-chip cooling (leaks, hot spots, pump failures, glycol degradation, biological growth) and the control strategy that keeps the loop within tolerance when the load swings by a factor of three between idle and peak inference. The paper is based on the cooling architecture HyperNext is deploying at Kakinada with our cooling-system partners.

The cooling architecture decision is not "air or water" any more. It is which liquid, in which loop, at what concentration, with what monitoring.

The physics of the crossover

Thermal carrying capacity of a fluid is the product of its mass flow rate, its specific heat capacity, and the temperature rise it tolerates. For air at standard conditions, specific heat is approximately 1.005 kJ per kg per kelvin. For pure water it is 4.18 kJ per kg per kelvin. Water at the same volumetric flow rate as air has approximately 830 times the mass. The combined effect: a water-based fluid carries roughly 3,500 times the heat per unit volume that air can, for the same temperature rise.

This is what makes direct liquid cooling practical at high density. A rack drawing 600 kW of heat needs to move approximately 30 cubic metres per second of air to reject that heat at a 15 degree Celsius temperature rise. The same heat can be rejected by 0.012 cubic metres per second of liquid coolant. About 13 litres per second through a manifold roughly the diameter of a fist. The physical engineering of moving 33 cubic metres per second of air through a single rack is physically impossible. Moving 13 litres per second of liquid is well-understood.

Why direct-to-chip, not immersion

There are two architectures for moving high-density heat with liquid. Direct-to-chip (D2C) brings cold-plate-based cooling to the GPU package, with coolant circulating through plates bolted to the silicon. Immersion cooling submerges the entire computing equipment in a dielectric fluid (single-phase mineral oil or two-phase fluorocarbon). Both work. Both are deployed in production by different operators.

HyperNext has standardised on D2C. The vendor base is broader. NVIDIA, AMD, HPE, and Hitachi Vantara all ship D2C-ready hardware as their primary high-density product line. Immersion cooling needs either OEM-supported immersion variants (limited availability) or aftermarket conversion (warranty complications, reliability complications, both). The operational disciplines are also closer to what data centre staff already know. Immersion needs different maintenance procedures, different safety protocols (especially for two-phase fluorocarbons), and different end-of-life equipment handling. The training investment is non-trivial.

That said, immersion has a real role for specific deployments. Especially for legacy hardware retrofit and for very-high-density edge deployments where the white space premium is decisive. We expect both architectures to coexist in the industry through the 2030s, with D2C as the volume mainstream and immersion as a specialty.

2. Glycol, not water

● Why glycol

The primary IT cooling loop in HyperNext facilities runs on a propylene glycol coolant mixture, not pure water. This is a closed-loop, makeup-free design in normal operation. The mixture is roughly 25 percent propylene glycol, 75 percent demineralised water, with corrosion inhibitor, biocide, and pH stabiliser additives. This is a deliberate engineering choice and it has cost. Pure water has 4.18 kJ per kg per kelvin specific heat. A 25 percent PG mixture has about 3.85 kJ per kg per kelvin. So a glycol loop needs roughly 8 percent more flow to carry the same heat. Pump power increases accordingly. The architecture is what allows the HyperNext published WUE target below 0.05 litres per kWh of IT load to be achievable.

Why we pay that cost anyway:

Pure water in a closed cooling loop is a biological reactor. Bacteria, algae, and biofilms colonise the system within weeks of commissioning and reduce thermal performance through fouling. Propylene glycol at 25 percent concentration suppresses biological growth in a way no water-only loop can sustain without aggressive biocide dosing that itself attacks loop materials.

Pure water is corrosive to the copper, brass, aluminium, and stainless steel mix found in modern cold plates and CDU heat exchangers. Galvanic action between dissimilar metals accelerates when the medium is uninhibited water. Propylene glycol coolant carries molybdate-based corrosion inhibitors that passivate metal surfaces and dramatically reduce corrosion rates.

Pure water needs continuous chemistry management to maintain conductivity below the threshold at which electrical leakage through the coolant becomes a concern. Glycol coolant maintains a stable low conductivity profile through its formulation, with the additives chosen to keep dielectric performance within IT-safe limits.

And: glycol coolant gives freeze protection. Most HyperNext campuses do not need freeze protection in the operational sense (interior India does not freeze) but the coolant inventory in chilled-water lines that pass through unconditioned plant rooms can drop close enough to dewpoint conditions during early-morning shoulder hours that ice formation in a stagnant section after a pump trip becomes a small but non-zero risk. A 25 percent PG mixture pushes the freeze point to roughly minus 10 degrees Celsius and removes that risk entirely.

Building secondary chilled-water loop is different. It runs ordinary chilled water (with low-glycol dosing for corrosion control) because it is fully enclosed within conditioned plant rooms and the heat exchange to the IT loop is through CDU brazed plate exchangers. The two-loop separation lets us pick the right chemistry for each role.

● Glycol chemistry monitoring

The IT coolant loop chemistry is monitored continuously on a 15-minute update cycle. The BMS tracks glycol concentration (target 25 percent +/- 2 percent), pH (target 8.5 to 9.5), conductivity (target below 50 microsiemens per centimetre), inhibitor reserve (molybdate concentration above 200 ppm), and particulate load (5 micron and finer filtration on bypass loops). Any out-of-spec parameter raises a BMS alarm and triggers the water-treatment vendor.

Annual chemistry cost runs approximately INR 6 lakh per megawatt of IT load. Glycol coolant replacement is on a four to five year cycle (full system flush and refill). Inhibitor top-up happens every six months. Biocide is dosed continuously through a side-stream pot feeder. The alternative is performance degradation that costs many times more in throughput lost over the life of the deployment.

3. The flow path

From chip to chiller

The heat path in a HyperNext D2C deployment runs through five engineered stages. Each has its own tolerances and its own failure modes.

Stage	What it is	Temperature range	Flow rate
1. Cold plate (silicon interface)	Copper plate bolted to GPU/CPU package	Tcase below 95 degrees C	0.4 L/s per chip
2. Rack manifold (PG loop)	Bulk supply/return to each cold plate	30 to 50 degrees C	13 L/s per rack
3. Coolant Distribution Unit (CDU)	Brazed plate exchanger PG to water	Hot 50 / Cold 30 degrees C	50 L/s per CDU (4 racks)
4. Building secondary loop (chilled water)	Chilled water from chiller plant to CDU	7 deg supply, 18 deg return	240 L/s per row
5. Chiller plant	Centrifugal chillers and condenser/dry cooler array	Variable	Centralised

Each stage has a temperature delta budget. Cold plate to coolant interface drops approximately 15 degrees Celsius. Cold plate to manifold delta is approximately 15 across the rack. CDU exchanger delta is approximately 12. Chiller delta is approximately 11. Total grid (chip to cold supply) is approximately 53 degrees Celsius. Room temperature does not appear in this chain. The cooling is sealed from room air.

Stages 1 to 3 run propylene glycol coolant in a closed loop. Stage 4 runs chilled water in a separate closed loop. Stage 5 is the building chiller plant. The CDU at stage 3 is the boundary between the IT-side glycol loop and the building-side water loop. Cross-contamination is prevented by the brazed plate construction of the exchanger and by differential pressure monitoring on either side.

The control loop

Flow rate and pump speed at each stage are independently controlled by the building management system. The primary control variable is GPU case temperature, reported to the BMS via the IT equipment own telemetry on a one-second update cycle. The control loop targets a case temperature of 75 to 85 degrees Celsius under all operating conditions, with hard limits at 95 (firmware-enforced) and the BMS attempting to keep operations well below the hard limit at all times.

The challenge is that load on the cooling system can change by a factor of three or more within seconds. A rack in inference idle draws approximately 200 kW. The same rack under aggressive inference workload draws 600 kW. The cooling system must respond to this without the case temperatures going outside their target band. This needs pump speed control on a sub-second cycle, fast-acting valves at the manifold level, and a CDU design with a thermal mass large enough to buffer the secondary loop against primary loop transients.

The HyperNext design uses variable-frequency-drive pumps on the secondary loop (controlled to maintain a fixed differential pressure across the CDUs) and electronically actuated control valves at the rack manifolds (controlled to maintain rack supply temperature within 1 degree Celsius of setpoint). The control algorithm is a cascaded PID with feedforward from the upstream IT load forecast, which the BMS produces from a one-minute-ahead prediction of expected workload. The feedforward term is what keeps the system stable through load transients. Without it the response would lag and case temperatures would briefly exceed targets.

4. The failure modes

Leaks

The most consequential failure mode of any liquid-cooled system is a leak. Coolant inside an energised data centre is a multi-hundred-thousand-rupee event at best, a multi-crore event at worst, and a tenant-trust-destroying event in the wrong combination of circumstances.

The HyperNext leak management strategy has three layers. The primary loop is designed for negative pressure. The PG loop runs below atmospheric pressure such that any breach draws air in rather than pushing coolant out. This is the dominant design feature that distinguishes the modern D2C systems from earlier liquid cooling. Every manifold connection, every cold-plate fitting, and every CDU joint has a leak detection sensor (capacitive rope or hot-wire) that signals to the BMS within seconds of any moisture detection. Isolation valves at the rack and row level allow the affected segment to be isolated within seconds of detection, with the surrounding racks continuing to operate on the redundant cooling path.

Of the three layers, negative-pressure design matters most. With it, a leak means a few drops absorbed into the rope sensor before isolation. Without it, a leak means coolant on the floor and electrical equipment in the line of fire. The pumping architecture to maintain negative pressure adds approximately 8 percent to the pump power budget. The engineering trade-off is worth it. Glycol is benign environmentally (propylene glycol is food-safe at the concentrations we use) but it is still a contamination risk to electrical equipment.

Hot spots

Even within a well-designed D2C system, individual cold plates can develop hot spots from poor manufacturing tolerance, from interface material degradation, or from particulate buildup in the manifold restricting flow to a specific plate. The BMS detects hot spots through rack-level temperature distribution. A single GPU running 5 degrees Celsius warmer than its neighbours is flagged for investigation.

Investigation is usually short. Re-seat the cold plate. Replace the thermal interface material. Flush the manifold. The work is done during a scheduled maintenance window with the affected rack drained and the redundant cooling path serving the room. Hot-spot frequency in the HyperNext Phase 1 deployment is approximately one event per rack per year, which is well within the maintenance budget envelope.

Pump failure

Pump failures in the primary PG loop and the secondary chilled-water loop are handled by N+1 redundancy at every pump stage. A single pump failure is invisible to the IT load. The redundant pump picks up the duty within seconds, the failed pump is isolated for replacement, and the system continues to operate at full capacity. Multiple simultaneous pump failures in the same loop are an outage scenario. The design target is mean-time-to-multiple-failure of greater than 50 years, which is the industry standard for Tier IV operation.

Glycol degradation

Propylene glycol coolant degrades over time. Oxidation produces organic acids that lower pH and consume the inhibitor reserve. Thermal cycling at the cold plate interface degrades the glycol polymer chain slowly. Particulates from incidental wear on pump impellers and valve seats accumulate in the loop. None of this is fast. None of it is undetectable. The chemistry monitoring described in Section 2 catches it before performance degradation becomes operationally significant. Full coolant replacement on a four to five year cycle resets the chemistry. Inhibitor top-ups every six months keep the reserve adequate between replacements.

Biological growth

Glycol at 25 percent concentration suppresses but does not eliminate biological growth. Some bacterial species tolerate propylene glycol at this concentration. Without biocide dosing those species can establish biofilms over time. The HyperNext loop carries a side-stream biocide feeder that maintains background biocide concentration above the kill threshold for the dominant species we have characterised in our water sources. Bypass-loop 5 micron filtration catches any biofilm fragments that detach from system surfaces. Annual biological audits (sampling at multiple points in the loop, culturing for the species of concern) verify that the biocide regime is working.

5. What we are still figuring out

● Two-phase D2C

The next architectural transition in liquid cooling is two-phase D2C. Instead of single-phase coolant that absorbs heat by warming up, the system uses a refrigerant that absorbs heat by changing state from liquid to vapour. Latent heat of vaporisation gives significantly more cooling capacity per unit mass than the sensible heat capacity of glycol coolant. Lower flow rates and smaller manifolds for the same heat removal.

Two-phase D2C is in production at a handful of hyperscalers for very-high-density deployments (1+ MW per rack) but the engineering is harder. The pressure dynamics of two-phase systems are more complex, the leak management is more critical (refrigerant leaks have environmental implications that glycol does not), and the supply chain for two-phase components is younger. HyperNext is monitoring the technology and expects to deploy it in the 2029 to 2030 horizon for capacity above the 600 kW per rack that Vera Rubin Ultra NVL576 establishes. For now, single-phase glycol coolant is the right call.

● Higher operating temperatures

Conservative target for cold plate supply temperature has been 25 to 30 degrees Celsius. The thermodynamic argument for raising it is strong. Higher supply temperature means smaller chiller plant, smaller delta-T across heat exchangers, and (above a threshold) ability to use ambient-air dry cooling without compressor work for much of the year. Some operators have deployed at 40 degrees Celsius cold plate supply. The trade-off is reduced thermal margin at the GPU. GPU vendors increasingly publish specifications that accept higher supply temperatures but the warranty position remains conservative.

The current HyperNext design targets 30 degrees Celsius cold plate supply with design flexibility to move to 35 as GPU vendor specifications mature. Energy savings from the higher supply temperature are approximately 4 to 7 percent of total cooling power.

● Heat reuse

The CDU return-side temperature in the HyperNext architecture is approximately 45 degrees Celsius. That heat is rejected to the building chilled-water loop and then to the atmosphere through the dry coolers. It is also above the temperature threshold at which it is usable for several industrial applications. The HyperNext architecture supports an optional heat reuse offtake: a tertiary loop teed off the chilled-water return that delivers 40 to 45 degrees Celsius water to co-located industrial users. The right offtake categories are district heating (limited applicability in India outside hill stations), greenhouse heating (relevant for fresh-produce operations co-located with data campuses), desalination pre-heating (relevant for coastal locations like Kakinada), and pre-heating for industrial processes that would otherwise heat from cold.

Heat reuse is offered as an option rather than a default because the offtake requires a customer at industrial scale within a few hundred metres of the campus. The economics work when that customer exists. Without that customer, the engineering is overhead. The HyperNext Kakinada masterplan reserves space for a heat reuse interface and includes the interface piping in the base build, so it can be activated when an industrial co-located customer signs up.

HEADLINES

- Air cooling stops working around 50 kW per rack. From there to the 600 kW per rack of Vera Rubin Ultra NVL576, everything is liquid.
- The HyperNext architecture uses single-phase propylene glycol coolant (25 percent PG, 75 percent demineralised water with corrosion inhibitor and biocide) in the primary IT loop. The secondary building loop is chilled water. The CDU heat exchanger is the boundary between them.
- Glycol gives us biological resistance, corrosion control, stable conductivity, and freeze protection in stagnant lines. The 8 percent pump power penalty is worth it.
- The five-stage flow path has approximately 53 degrees Celsius of total delta budget. Each stage contribution is engineered separately and the control loop holds case temperatures within 1 degree Celsius of setpoint through load transients.
- Failure modes that matter: leaks (handled by negative pressure design and isolation valves), hot spots (rack-level temperature monitoring), pump failures (N+1 redundancy), glycol degradation (chemistry monitoring and 4 to 5 year replacement cycle), biological growth (biocide dosing and side-stream filtration).
- Next architectural transitions are two-phase D2C (likely 2029 to 2030) and higher cold plate supply temperatures (incremental, dependent on GPU vendor specifications).

6. Glycol coolant chemistry deep dive

Section 2 introduced the propylene glycol coolant mixture HyperNext uses in the primary IT loop. The section below covers the chemistry in operational detail. It is intended for the cooling engineer, the chemistry monitoring team, and the audit reviewer.

● Mixture composition

Component	Concentration	Function
Propylene glycol (USP-grade)	25% v/v nominal, 23-27% acceptable range	Freeze protection, biological growth suppression, hydrocarbon-compatible
Demineralised water	74% v/v nominal	Primary heat carrier, low conductivity baseline
Sodium molybdate (corrosion inhibitor)	250 ppm target, 200 ppm minimum reserve	Passivates ferrous and copper surfaces, prevents galvanic corrosion
Sodium nitrite (supplementary inhibitor)	100 ppm target	Reinforces ferrous passivation, kept below 200 ppm to avoid copper attack
Sodium tetraborate (pH buffer)	500 ppm target	Maintains pH in 8.5-9.5 range against organic acid formation
Tolyltriazole (copper-specific inhibitor)	20 ppm target	Forms protective film on copper alloy surfaces
Glutaraldehyde (biocide)	50 ppm running concentration via side-stream dosing	Suppresses bacterial growth, including PG-tolerant species
Sodium hydroxide (pH adjustment)	As needed to maintain pH target	Compensates for inhibitor consumption

● Why these specific additives

Sodium molybdate as the primary corrosion inhibitor is chosen over the traditional alternative (sodium nitrite) because molybdate gives equivalent ferrous protection without the toxicity and disposal complications of nitrite-dominant systems. The IT cooling loop will eventually be drained for replacement (4 to 5 year cycle) and a molybdate-dominant formulation produces a waste stream that is easier to dispose under Indian hazardous waste regulations.

Sodium tetraborate as the pH buffer is chosen because the natural pH drift in a closed glycol loop is acidic (organic acid formation from PG oxidation). The buffer keeps the loop in the slightly alkaline range where most metals are passive. Without buffering, the loop would drift toward pH 6 to 7 within months, accelerating copper corrosion.

Tolyltriazole specifically targets copper. The cold plates in modern direct-to-chip systems are copper or copper-alloy with high-purity copper for the chip-contact surface. Tolyltriazole forms a monolayer film on copper that protects against electrochemical attack while not interfering with thermal performance. The 20 ppm running concentration is conservative; vendor specifications often allow up to 50 ppm.

Glutaraldehyde as the running biocide is chosen because propylene glycol at 25 percent suppresses but does not eliminate biological growth. Several bacterial species tolerate PG at this concentration and can establish biofilms over time. Glutaraldehyde at 50 ppm running concentration kills these species without attacking loop materials. Alternative biocides (isothiazolones, bromine releasers) were evaluated but glutaraldehyde gives the best material compatibility profile for the HyperNext cold plate metallurgy.

Monitoring frequency and parameters

```
CONTINUOUS MONITORING (real-time to BMS)
  Loop pressure, supply and return    1-second
  Loop temperature, supply and return  1-second
  Loop flow rate per CDU              1-second
  Leak detection rope status           1-second
  Reservoir level                      1-second
  Conductivity                        15-second

SCHEDULED CHEMISTRY MONITORING
  Glycol concentration (refractometer)  15-minute
  pH                                     15-minute
  Inhibitor reserve (molybdate)          Daily
  Inhibitor reserve (nitrite)            Daily
  Tolyltriazole concentration            Weekly (titration)
  Biocide concentration                  Weekly (titration)
  Total dissolved solids                 Weekly
  Bacterial count (heterotrophic plate)  Monthly
  Particulate count (5 micron)           Weekly

CALIBRATION AND VERIFICATION
  Sensor calibration                     6-monthly
  Independent lab sample comparison      Quarterly
  Full chemistry audit                   Annually
```

What triggers operator action

The BMS chemistry alarm rules are tiered. Parameters within their nominal range produce no alarm. Parameters in the warning band trigger an INFO alarm to the chemistry team for review within the next shift. Parameters in the warning band for more than 48 hours escalate to MAJOR. Parameters outside the warning band trigger an immediate MAJOR alarm. Parameters at safety-critical extremes (e.g., conductivity above 200 microsiemens per centimetre, which could allow electrical leakage through the coolant) trigger a CRITICAL alarm and may initiate automatic CDU isolation under the most extreme conditions.

Parameter	Nominal	Warning band	Major alarm	Critical
PG concentration	23-27 percent	20-23 or 27-30	below 20 or above 30	—
pH	8.5-9.5	8.0-8.5 or 9.5-10.0	below 8.0 or above 10.0	below 7.0
Conductivity (microsiemens/cm)	below 50	50-100	100-200	above 200
Molybdate reserve (ppm)	above 200	150-200	100-150	below 100
Biocide concentration (ppm)	40-60	20-40 or 60-100	below 20 or above 100	—
Bacterial count (CFU/ml)	below 100	100-1,000	1,000-10,000	above 10,000

7. Heat transfer calculations

The five-stage flow path in Section 3 has a temperature delta budget of approximately 53 degrees Celsius. The section below shows how the engineering arrives at that budget and where the trade-offs are.

● Stage 1: chip to cold plate

The thermal resistance from chip junction to cold plate fluid is the sum of three resistances in series. Junction to package case (set by the silicon design and the package thermal interface, around 0.04 degrees Celsius per watt for current-generation GPUs). Case to cold plate (set by the thermal interface material, typically 0.02 to 0.05 K/W with high-grade indium or silver-loaded TIMs). Cold plate to fluid (set by the cold plate internal geometry and the fluid-side heat transfer coefficient).

```
EXAMPLE: 1.2 kW GPU CHIP
  Junction temperature target              85 deg C (margin to 95 deg C limit)
  Junction to case thermal resistance      0.04 K/W (vendor spec)
  Case to TIM thermal resistance           0.03 K/W (high-grade TIM)
  TIM to cold plate resistance             0.05 K/W
  Cold plate to fluid resistance           0.012 K/W (turbulent flow, fin geometry)

  Total thermal resistance                 0.132 K/W
  Heat flow at full TDP                    1200 W
  Temperature rise junction to fluid       158 deg C  (UNACCEPTABLE)
```

The arithmetic above does not work. A 1.2 kW chip cannot maintain 85 deg C junction with a total resistance of 0.132 K/W. The cold plate to fluid resistance has to be much lower. This is achieved by spreading the chip footprint across a larger fin area inside the cold plate (effective heat transfer area roughly 4 to 6 times the chip footprint) and increasing the fluid velocity at the fin surface.

```
REVISED COLD PLATE DESIGN
  Effective fluid-side area                4.5 x chip footprint
  Reynolds number at design flow           18,000 (turbulent regime)
  Local heat transfer coefficient          12,000 W/m^2-K
  Effective cold plate to fluid resistance 0.025 K/W (acceptable)

  Total thermal resistance                 0.145 K/W -> revised to 0.075 K/W with paralle
  Temperature rise at full TDP             90 K
  Junction at fluid 25 deg C               115 deg C  (still above limit)
```

The remaining gap is closed by lowering the fluid inlet temperature. At 25 deg C the system does not work. At 30 deg C inlet, with the revised cold plate, the junction sits at approximately 120 deg C — still over limit.

The HyperNext design specifies 32 deg C cold plate fluid supply maximum, which combined with the cold plate design above gives a worst-case junction of approximately 85 to 90 deg C.

The arithmetic illustrates why direct-to-chip cooling is engineering at the edge. Small changes in cold plate fluid temperature have large effects on chip junction temperature. The control system has to maintain inlet temperature within plus or minus 1 deg C of setpoint to keep junction temperatures within the design band.

Stage 2: cold plate to rack manifold

The rack manifold collects coolant from all the cold plates in the rack and returns it to the CDU. The temperature rise across the manifold is essentially the average of the individual cold plate temperature rises, weighted by flow.

```
RACK MANIFOLD HEAT BUDGET
  Rack heat load                           600 kW
  Coolant mass flow per rack               12 kg/s (PG mixture, density 1,015 kg/m^3)
  Specific heat capacity (25% PG)          3,850 J/kg-K

  Bulk temperature rise across the rack    delta_T = Q / (m_dot * Cp)
                                           = 600,000 / (12 * 3,850)
                                           = 13.0 deg C
```

The bulk fluid rises about 13 degrees from supply to return at the rack level. The design budget allows up to 15 degrees, leaving 2 degrees of headroom for non-uniform heat distribution and partial-flow fault conditions.

Stage 3: CDU heat exchanger

The CDU is a brazed plate heat exchanger that transfers heat from the primary glycol loop to the secondary chilled-water loop. The exchanger sizing balances three constraints. The temperature pinch (the smallest temperature difference between the two streams at any point in the exchanger), the pressure drop on each side (sets pump power consumption), and the physical size and cost.

```
CDU SIZING FOR A 4-RACK GROUP (2.64 MW heat load)
  Primary side (PG mixture)
    Inlet 45 deg C, outlet 32 deg C        delta_T = 13 deg C
    Flow rate                               2.64 MW / (3.85 kJ/kg-K * 13 K) = 52.6 kg/s

  Secondary side (chilled water)
    Inlet 7 deg C, outlet 18 deg C          delta_T = 11 deg C
    Flow rate                               2.64 MW / (4.18 kJ/kg-K * 11 K) = 57.4 kg/s

  Log mean temperature difference          [(45-18) - (32-7)] / ln[(45-18)/(32-7)] = 26.0
```

```
    Required heat transfer coefficient (U)   2,500 W/m^2-K (typical for brazed plate)
    Required area (Q / U / LMTD)             40.6 m^2

    Selected exchanger                       APV B25 series, 45 sq m, brazed plate
    Margin                                   11% on heat transfer area
```

The CDU is the engineering boundary between the glycol IT loop and the building chilled-water loop. The brazed plate construction means that even in the event of a plate failure, the two fluids do not mix and there is no path for glycol to contaminate the building chilled-water system or vice versa. Pressure differential monitoring on both sides catches any plate degradation early.

8. Pump and pipe sizing methodology

Pump and pipe sizing for the IT-side glycol loop is the design problem that determines whether the cooling architecture is robust under transient load conditions. The section walks through the methodology.

Primary loop pump sizing

The primary loop pump has to overcome the friction losses through the cold plates, the rack manifolds, the supply and return piping, the CDU heat exchanger primary side, and the strainer and filter assemblies. Each of these contributes a pressure drop that the pump must supply against.

```
PRESSURE DROP BUDGET PER CDU (4 racks, 52.6 kg/s)
  Cold plates (manufacturer-rated)         1.2 bar
  Rack manifolds                            0.4 bar
  Rack-to-CDU piping (40 m, 6-inch)         0.3 bar
  CDU primary side                          0.6 bar
  Strainer and filter                       0.3 bar
  Margin (15%)                              0.4 bar
                                            ---------
  Total system head                         3.2 bar (32 m of fluid column equivalent)

  Pump duty point                           52.6 kg/s at 32 m head
  Selected pump                             Grundfos NK 80-200 or equivalent
  Motor rating                              22 kW
  Pump efficiency at duty point             82%
  Hydraulic power                           17.6 kW

  N+1 redundancy: 2 pumps installed, 1 carrying load, 1 standby
  Both pumps controlled to maintain differential pressure across the CDUs (PID)
```

The pump rating is sized for the worst-case operating condition (peak flow at peak inlet temperature, with all filters at maximum allowed pressure drop). Normal operating conditions use about 60 to 70 percent of pump rated power. Variable frequency drives let the pump speed reduce when full flow is not required, saving roughly 30 percent of pump energy under typical operating conditions.

Pipe sizing

The supply and return piping from the CDU to the racks is sized for two constraints. The pressure drop budget (above) and the maximum acceptable fluid velocity (to avoid erosion of the inner pipe wall over the asset lifetime). For PG mixture in carbon steel pipe, the recommended maximum velocity is 2.5 m/s. Sustained velocities above 3 m/s cause measurable erosion at elbows and reducers within 10 to 15 years.

```
PIPE SIZING FOR THE 52.6 KG/S DUTY
  Volumetric flow rate                      52.6 / 1015 = 0.0518 m^3/s
  Target velocity                           2.0 m/s
  Required cross-sectional area             0.0518 / 2.0 = 0.0259 m^2
  Required internal diameter                182 mm

  Standard pipe selection                   8-inch nominal (203 mm internal)
  Actual velocity at duty                   1.61 m/s
  Margin against erosion limit              36%
```

The 8-inch nominal pipe gives margin against erosion and provides headroom for future flow increase if rack power per CDU group grows. Sizing more tightly to the current duty point creates rework risk if the next generation of GPUs increases the per-rack heat load.

Pipe materials

The HyperNext design uses carbon steel for the bulk of the primary loop piping, with stainless steel at the CDU connection points and the cold plate manifolds. The carbon steel is acceptable for the PG-water mixture because the corrosion inhibitor regime in Section 6 passivates the inner surface. Without proper inhibitor maintenance, carbon steel would corrode rapidly. With it, the inner surface is stable for 20+ year operating life.

The piping is welded throughout the bulk runs. Flanged connections are limited to equipment terminations (pumps, CDU, strainers) where eventual removal for maintenance is expected. Threaded connections are not permitted anywhere in the primary loop because threads create stress concentrations and leak paths under thermal cycling.

9. References and standards

- **ASHRAE TC 9.9.** Liquid Cooling Guidelines for Datacom Equipment Centers, 3rd edition, 2023. The reference for facility liquid cooling design.
- **ASHRAE TC 9.9.** Water-cooled Servers, second edition. The reference for direct-to-chip cooling design.
- **Open Compute Project.** Liquid Cooling Specification, current revision. The OCP framework for rack-level liquid cooling.
- **ASTM D1384.** Standard Test Method for Corrosion Test for Engine Coolants in Glassware. The corrosion test methodology for inhibitor evaluation.
- **ASTM D2570.** Standard Test Method for Simulated Service Corrosion Testing of Engine Coolants. Used for inhibitor lifetime characterisation.
- **ASTM D3306.** Standard Specification for Glycol Base Engine Coolant. The reference for propylene glycol specification.
- **BS EN 13753.** Heat exchangers. Specifications for brazed plate heat exchangers. The reference for CDU sizing.
- **Crane Technical Paper 410.** Flow of Fluids Through Valves, Fittings, and Pipe. The reference for pressure drop calculations.
- **NVIDIA.** Direct Liquid Cooling Implementation Guide for HGX platforms. The vendor specification for cold plate flow and temperature requirements.
- **HyperNext Research HN-RP-002.** 800VDC Power Architecture. The companion paper covering the power side of the same architectural transition.

HyperNext Research

We publish engineering and policy papers because the Indian conversation about AI infrastructure needs more substance than marketing material provides. The papers state methodology openly so other operators can run the same analysis on their own facilities. They report findings that may not flatter the HyperNext commercial position. They get review from the engineering team and the communications partners.

Correspondence on methods, figures, and conclusions: **hello@hypernxt.com**. We read every email.

HN-RP-006 · Liquid Cooling at 600 kW per Rack
29 April 2026 · v1.0

www.hypernxt.com/research
hello@hypernxt.com · +91 99784 23333