Closed-loop water systems in data centers are cooling systems that continuously recirculate the same water rather than consuming fresh water in a "once-through" process. They are becoming more important as AI workloads increase rack power density and as operators face pressure to reduce water usage and improve sustainability.

Why data centers need advanced cooling

Modern data centers convert most electrical power into heat. Traditional air cooling struggles once rack densities climb into the 30–100+ kW range common in AI and high-performance computing clusters. Water is much more efficient at carrying heat than air, so operators increasingly use liquid-based systems.

The challenge is balancing:

- Heat removal
- Energy efficiency
- Water consumption
- Reliability
- Cost

Closed-loop systems address several of these simultaneously.

What "closed loop" means

In a closed-loop cooling system:

1. Water (or a coolant mixture) circulates through pipes.
2. It absorbs heat from servers or cooling equipment.
3. The warmed liquid transfers heat elsewhere through a heat exchanger.
4. The same liquid is cooled and reused repeatedly.

Very little water is lost during normal operation.

This contrasts with:

- **Open-loop systems**, which continuously consume or evaporate water
- **Once-through cooling**, where water is drawn from a source and discharged after use

---

Main types of closed-loop cooling in data centers

1. Chilled water loops

This is the classic enterprise data center approach.

How it works

- Chillers cool water
- Pumps circulate chilled water through cooling coils
- Air handlers cool server-room air
- Heat returns to the chiller plant

Characteristics

- Mature and reliable
- Works well for conventional racks
- Often paired with economizers for efficiency

Limitation

Air is still the final cooling medium at the rack level, which becomes inefficient at very high densities.

---

2. Direct-to-chip liquid cooling

This is becoming common in AI infrastructure.

How it works

Cold plates are attached directly to CPUs and GPUs. Coolant flows through the plates and removes heat at the source.

Benefits

- Handles very high thermal loads
- Reduces fan energy
- Supports dense AI racks

Typical loop structure

There are often multiple loops:

- Facility water loop
- Secondary coolant loop
- Rack manifold loop

Heat exchangers isolate them so server coolant never mixes with facility water.

Companies like NVIDIA and Supermicro actively support these architectures in modern AI clusters.

---

3. Immersion cooling

Servers are submerged in nonconductive liquid.

Two major forms

Single-phase immersion

The fluid stays liquid while circulating heat away.

Two-phase immersion

The liquid boils on hot components, vapor rises, condenses, and returns.

Advantages

- Extremely high cooling efficiency
- Very dense compute possible
- Minimal airflow infrastructure

Challenges

- Operational complexity
- Hardware compatibility
- Maintenance procedures

Companies such as Submer and LiquidStack specialize in these systems.

Water usage effectiveness (WUE)

A major metric for closed-loop systems is Water Usage Effectiveness.

$$WUE = \frac{\text{Annual Water Usage}}{\text{IT Equipment Energy}}$$

Lower WUE means less water consumed per unit of compute.

Traditional evaporative cooling can consume millions of gallons annually. Some closed-loop systems reduce operational water use dramatically because they:

- Avoid evaporation
- Reuse coolant continuously
- Use dry coolers instead of cooling towers

Key components in a closed-loop system

Heat exchangers

Transfer heat between loops without mixing fluids.

Pumps

Maintain coolant circulation.

Coolant distribution units (CDUs)

Control temperature, pressure, and flow to server racks.

Dry coolers

Reject heat to outdoor air without evaporating water.

Expansion tanks

Handle thermal expansion of the coolant.

Water treatment systems

Prevent:

- Corrosion
- Biological growth
- Mineral scaling

Advantages

Water conservation

Especially important in drought-prone regions.

Better efficiency

Liquid cooling transfers heat more effectively than air.

Higher rack density

Supports modern AI and HPC deployments.

Lower fan energy

Servers require less airflow.

More stable thermal control

Coolant temperatures can be tightly managed.

Drawbacks and engineering challenges

Higher upfront cost

Liquid infrastructure is expensive.

Leak management

Even well-designed systems require detection and containment.

Maintenance complexity

Fluid chemistry and pump systems require specialized expertise.

Retrofit difficulty

Older air-cooled facilities may need major redesigns.

Heat rejection still matters

Even closed loops ultimately must dump heat somewhere:

- Outdoor air
- District heating
- Cooling towers
- Nearby industrial processes

Emerging trends

Warm-water cooling

Systems run at higher coolant temperatures, reducing chiller use.

Heat reuse

Some facilities redirect waste heat to:

- Buildings
- Greenhouses
- District heating systems

For example, operators in parts of Finland and Denmark increasingly integrate data center heat into municipal heating networks.

AI-driven thermal optimization

Machine learning is being used to dynamically adjust:

- Pump speeds
- Flow rates
- Workload placement
- Cooling setpoints

Hybrid systems

Facilities may combine:

- Air cooling
- Direct liquid cooling
- Rear-door heat exchangers
- Immersion systems

Closed-loop vs open-loop at a glance

Feature	Closed-loop	Open-loop
Water reuse	High	Low
Water consumption	Low	High
Complexity	Higher	Lower
Sustainability	Better	Worse
AI workload suitability	Excellent	Moderate
Maintenance	More specialized	Simpler

Why this matters now

The rise of generative AI has dramatically increased power density in data centers. AI clusters using GPUs from companies like AMD and NVIDIA can push rack power far beyond what traditional air cooling can efficiently support.

As a result, many hyperscalers and colocation providers are redesigning facilities around closed-loop liquid cooling architectures to:

- Reduce water usage
- Improve energy efficiency

- Support denser compute
- Meet sustainability targets
- Prepare for future AI hardware generations

Even in a "closed-loop" data center cooling system, cooling towers can still consume significant water if the system rejects heat through evaporation.

The important distinction is:

- **Closed-loop coolant circuit** → the same water/coolant recirculates through the servers and heat exchangers
- **Cooling tower loop** → may still evaporate water to dump heat outdoors

So a data center can have a closed-loop IT cooling system but still use water at the cooling tower.

Where the water goes

Cooling towers mainly lose water through:

1. Evaporation (largest loss)

Heat is removed when a small portion of water evaporates into the air.

Rule of thumb:

$$\text{Evaporation Loss} \approx 0.00085 \times \text{Flow Rate} \times \Delta T$$

A simpler engineering estimate:

About **1.8 gallons of water evaporate per ton-hour of cooling**.

(1 refrigeration ton = 12,000 BTU/hr.)

---

2. Blowdown

Some water is intentionally drained to remove:

- Minerals
- Salts

- Contaminants

Without blowdown, scaling and corrosion become severe.

Typical blowdown adds another:

- ~5–20% on top of evaporation losses

---

3. Drift

Tiny droplets escape the tower airflow.

Modern drift eliminators make this relatively small.

Typical data center water consumption

A medium-to-large hyperscale data center may consume:

- **Hundreds of thousands to several million gallons per day**
- Depending on:
  - Climate
  - Cooling design
  - IT load
  - Humidity
  - Water reuse systems

For example:

Facility Type	Approximate Water Use
Traditional evaporative-cooled hyperscale	1–5+ million gallons/day
Closed-loop with cooling towers	Lower, but still substantial
Closed-loop with dry coolers	Near-zero operational water use

Example calculation

Suppose a data center rejects:

- 50 MW of heat

That is roughly:

$$50\text{ MW} \approx 14{,}200 \text{ refrigeration tons}$$

Using the common rule:

$$1.8 \times 14{,}200 \approx 25{,}500 \text{ gallons/hour}$$

Over a full day:

$$25{,}500 \times 24 \approx 612{,}000 \text{ gallons/day}$$

That's just evaporation, before blowdown losses.

A very large AI-focused campus could exceed several million gallons/day in hot weather.

Why AI data centers increase water demand

AI clusters have:

- Much higher rack densities
- Larger thermal loads
- More continuous utilization

A conventional enterprise rack might use:

- 5–15 kW

Modern AI racks can exceed:

- 80–150+ kW

More heat means more cooling demand and potentially more evaporation if cooling towers are used.

How operators reduce tower water use

Dry coolers

Use outdoor air instead of evaporation.

Pros:

- Near-zero water use

Cons:

- Less efficient in hot climates
- Higher electricity use

Adiabatic cooling

Only sprays water during hot conditions.

Cuts annual water use substantially.

Warm-water liquid cooling

Allows higher coolant temperatures, reducing or eliminating chillers.

Water reuse

Some facilities use:

- Reclaimed wastewater
- Gray water
- Treated industrial water

Instead of potable drinking water.

Hybrid systems

Switch dynamically between:

- Dry cooling
- Evaporative cooling

Based on weather conditions.

Important nuance

People often say "closed-loop systems use little water," but that is only fully true if the heat rejection side is also closed and non-evaporative.

There are really two separate questions:

1. Is the IT cooling loop closed?
2. Does the heat rejection system evaporate water?

A facility can have:

- Closed-loop liquid cooling internally
- But still consume massive amounts of water externally through cooling towers

That's one of the biggest sustainability debates around modern AI infrastructure.

There are already 100 MW–scale data center campuses designed to operate with little or no cooling tower usage, and that trend is accelerating because of water constraints and AI growth.

However, there's an important distinction:

- **No cooling towers at all** → possible, but technically harder and usually less energy efficient in hot climates
- **Minimal cooling tower use** → much more common
- **Seasonal or hybrid tower use** → currently the dominant approach for very large AI campuses

The main alternatives to cooling towers

1. Dry cooling (air-cooled heat rejection)

Instead of evaporating water, the facility uses giant radiator-like systems outdoors.

How it works

Warm water from the data center passes through finned coils while fans blow outdoor air across them.

Very similar to:

- Automotive radiators
- Power plant dry coolers

Water use

Operational water use can approach near-zero.

Downsides

The hotter the outside air gets, the less effective dry cooling becomes.

That creates:

- Higher electricity consumption
- Larger equipment footprints
- Reduced cooling efficiency during heat waves

Why 100 MW is difficult without towers

A 100 MW AI campus is rejecting roughly:

$$100\text{ MW} \approx 341{,}000{,}000\text{ BTU/hr}$$

That is an enormous amount of heat.

Without evaporation, the only way to reject it is:

- Huge airflow volumes

- Very large heat exchangers
- Higher coolant temperatures

So tower-free systems tend to require:

- Massive dry cooler farms
- Large fan energy budgets
- Advanced liquid cooling architectures

Places where tower-free AI campuses make the most sense

Cool climates

Examples:

- Sweden
- Norway
- Finland
- Northern U.S. states
- High-altitude regions

Cold outdoor air dramatically improves dry cooling performance.

Regions with severe water constraints

Examples:

- Parts of Arizona
- Parts of Nevada
- Middle East projects

Water availability can become more limiting than power availability.

Architectures enabling no-tower designs

Direct-to-chip liquid cooling

Warm-water liquid cooling allows coolant temperatures high enough for dry coolers to work efficiently.

Instead of:

- 45–55°F chilled water

Some modern systems operate at:

- 80–100°F coolant temperatures

That can eliminate chillers entirely in some climates.

Rear-door heat exchangers

Heat is removed directly at the rack, reducing room-level cooling demand.

Immersion cooling

Immersion systems can run at relatively high fluid temperatures, making dry rejection easier.

Companies like Microsoft have explored immersion and advanced liquid cooling for dense AI infrastructure.

Real-world examples and trends

Meta

Meta has emphasized water-efficient designs and has built facilities using significant air-side economization and reduced-water strategies.

Microsoft

Microsoft has announced data center designs that can operate without water for cooling during large parts of the year and has developed closed-loop liquid systems intended to reduce evaporative water use substantially.

Google

Google has pursued "water-positive" goals and uses climate-specific cooling strategies, including reduced-water and reclaimed-water approaches.

NVIDIA ecosystem deployments

Many new GPU superclusters are being designed around liquid cooling plus dry coolers because:

- AI rack densities are exploding
- Water permits are getting harder to obtain
- Communities increasingly oppose large potable-water consumption

---

The tradeoff triangle

Operators are balancing three competing goals:

Goal	Easier With
Lowest energy use	Evaporative towers
Lowest water use	Dry cooling
Lowest capital cost	Traditional chilled water

You usually cannot optimize all three simultaneously.

What's becoming common at 100 MW scale

The most likely near-term architecture is:

Hybrid cooling

Most large AI campuses are moving toward:

- Closed-loop liquid cooling internally
- Dry cooling most of the year
- Limited evaporative assist during extreme heat

This dramatically cuts water use without requiring oversized dry systems.

Fully towerless 100+ MW campuses

These are technically feasible today, especially in:

- Cooler climates
- Locations with cheap renewable power
- Regions where water is politically constrained

But they generally require:

- Higher electrical overhead
- Larger heat rejection infrastructure
- More advanced thermal engineering
- Higher upfront capital cost

As AI densities rise, many experts expect:

- More warm-water liquid cooling
- More immersion cooling
- More dry-cooled campuses
- Less dependence on evaporative towers over time

especially in water-stressed regions of the U.S. and Europe.

A fully closed-loop, dry-air-cooled data center (meaning no evaporative cooling towers and little or no operational water use) solves the water problem, but it introduces major engineering and economic tradeoffs.

At hyperscale AI campus sizes — especially 100 MW+ — the disadvantages become significant.

Why AI changes the equation

Traditional cloud data centers already consumed large amounts of power, but AI dramatically increases intensity.

A modern AI campus may require:

- 100–500+ MW

For comparison:

- **100 MW can power tens of thousands of homes**
- **500 MW approaches utility-scale industrial demand**

Some proposed AI campuses now rival:

- Aluminum smelters
- Heavy industrial complexes
- Small cities

The grid problem is not just energy — it's timing

The biggest issue is often peak demand.

AI clusters may operate:

- 24/7
- Near full utilization
- With little flexibility

That stresses:

- Transmission systems
- Generation reserves
- Grid balancing operations

Unlike some industrial loads, AI compute demand is often economically incentivized to stay continuously active.

Regional examples

Virginia

Northern Virginia hosts the world's largest concentration of data centers.

Utilities there have warned that:

- Grid expansion is struggling to keep pace
- Transmission buildouts are accelerating
- Electricity demand forecasts have surged because of AI

Georgia

Georgia has become a major data center target because of:

- Available land
- Utility infrastructure
- Favorable regulation

But rapid growth has also raised debates over:

- Ratepayer impacts
- Power generation expansion
- Water use
- Long-term grid planning

Texas

Texas attracts hyperscalers because of:

- Cheap power
- Deregulated markets
- Renewable growth

But large loads can also intensify:

- Transmission congestion
- Reliability concerns during extreme weather

The broader infrastructure issue

Data centers are often marketed as “clean digital infrastructure,” but physically they function more like:

- Industrial campuses
- Utility complexes
- Heavy infrastructure nodes

especially at modern AI scales.

Their environmental footprint increasingly includes:

- Power generation demand
- Water demand
- Land transformation
- Transmission expansion
- Construction-related erosion

which is why many regions are rethinking zoning and permitting for hyperscale AI development.