



AI推理多元化液冷实践分享

田锋

超云 互联网与AIDC业务总经理

AI推理多元化液冷实践

田锋

超云 互联网与AIDC业务总经理

成立于2010年，由CEC和云基地联合注资，国家高新技术企业，国家级“专精特新”企业。

产品覆盖加速计算、信创整机、通用计算及私有云等领域。

秉承“让数据中心更简单”的价值主张，超云坚持自主创新、数实融合的发展路线，超云践行“生态共荣”及“AI IN ALL”战略，积极参与整合上下游产业链，助力千行百业实现数字化和智能化。



国内Tier1 AI推理设备制造商



NO.1

2023~2025年
X86服务器增速

NO.1

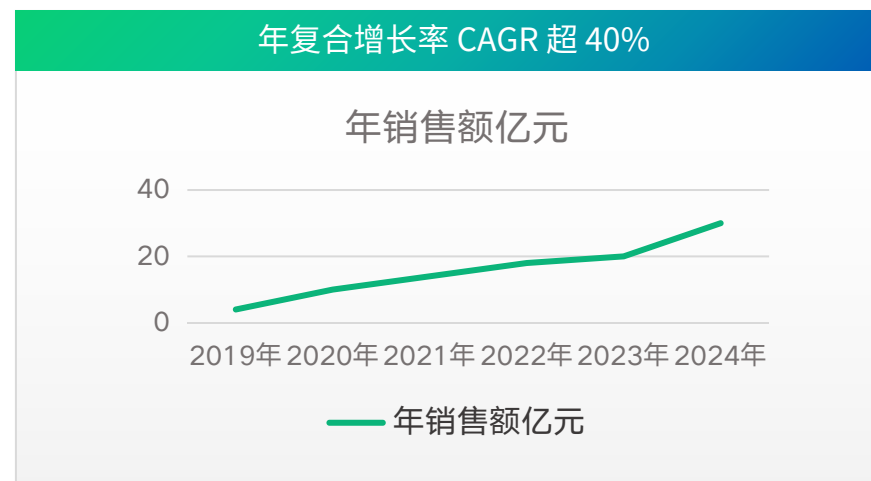
推理业务
总体增速

TOP5

液冷服务器
总体排名

TOP5

信创整机
总体排名



超过7年液冷经验

2020-2023

低碳数据中心先行者

- 丰富多元产品类型：通用机型、GPU机型、整机柜液冷、液冷微模块、集装箱液冷
- 液冷标准核心成员
- 浸没预研

2019

液冷元年

- 启动研发投入
- 国内最早一批布局液冷研发
- 绿色高效数据中心建设提供的最优产品解决方案

2024-2025

上榜液冷IDC

- Intel联合实验室
- 浸没液冷产品化
- 全栈式液冷解决方案
- 互联网浸没案例落地
- 热控联合实验室和人才培养基地

2026-2028

液冷AI算力首选品牌

- 全栈液冷产品能力
- 差异化技术路线
- 全生命周期服务

液冷专利筑基，标准领航




- GB/T 《信息技术 服务器及存储设备用液冷部件技术规范》
- T/CESA 1271-2023 《液冷式高性能计算机 技术参考架构》
- T/CESA 1272-2023 《高性能计算机 浸没式液冷技术要求》
- T/CESA 2443-2025 《冷板式液冷整机柜服务器 技术要求》

部署液冷专利 20+

液冷国标 1项 | 团标 3项

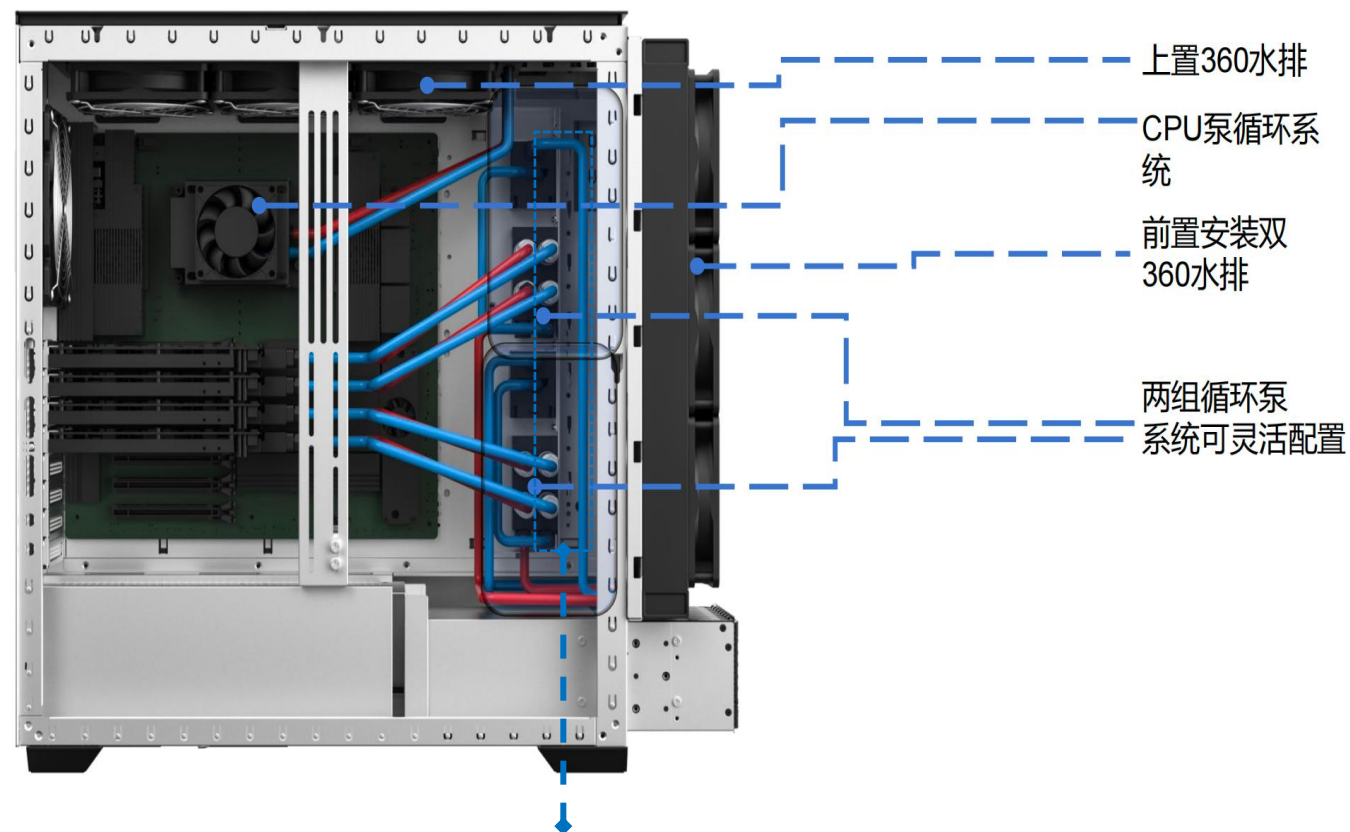
全栈液冷设备



	2024				2025				2026				
	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	Q1	Q2	Q3	Q4	
通用	 EGS				 BHS				 OKS				
AI	 4U8GPU&OAM EGS				 4U16GPU EGS				 6U8GPU BHS		 6U 16GPU BHS		 液冷工作站 4GPU
整机柜	 19英寸标准机柜 支持二总线液冷/48VDC 单柜120kW+				 19英寸标准机柜 水/电/网三总线 单柜120kW+								
解决方案	 液冷微模块				 冷板液冷预制化 方案				 浸没液冷预制化 方案				
浸没液冷					 Superrack-200				 7U PCIe浸没液冷产品				

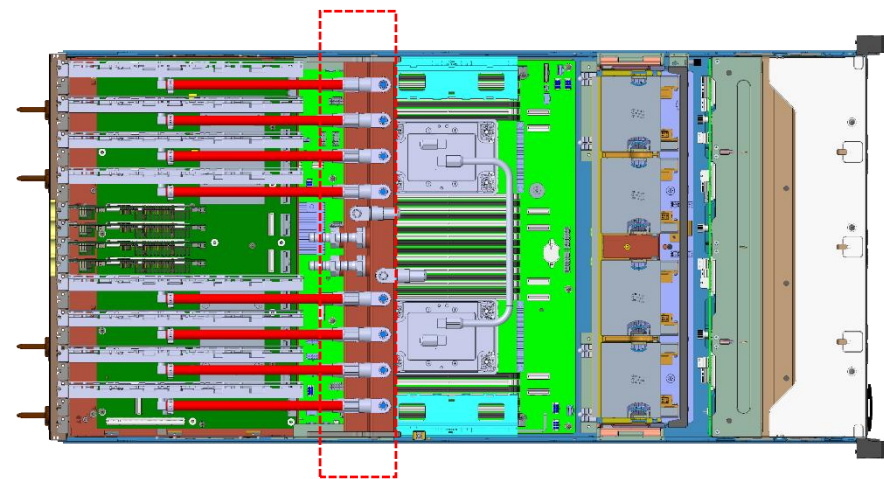
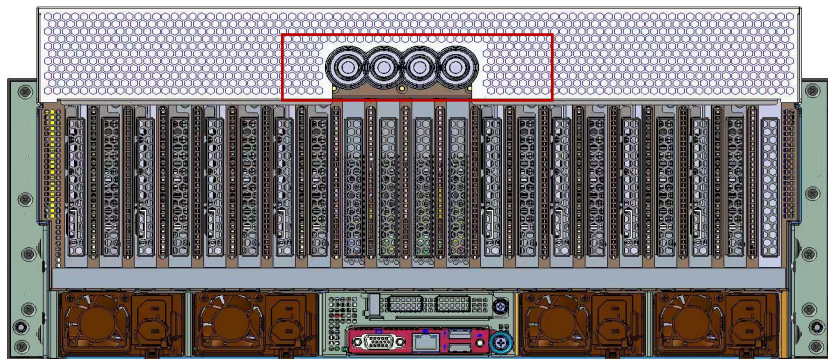
70B-超云CloseLoop冷板液冷四卡工作站

OpenClaw专版



QD, 无滴漏运维

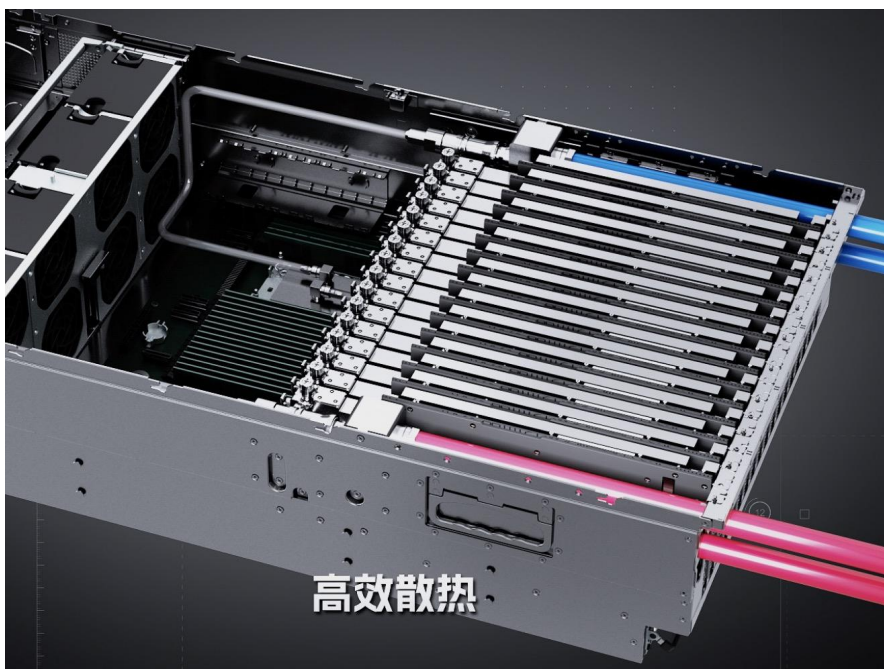
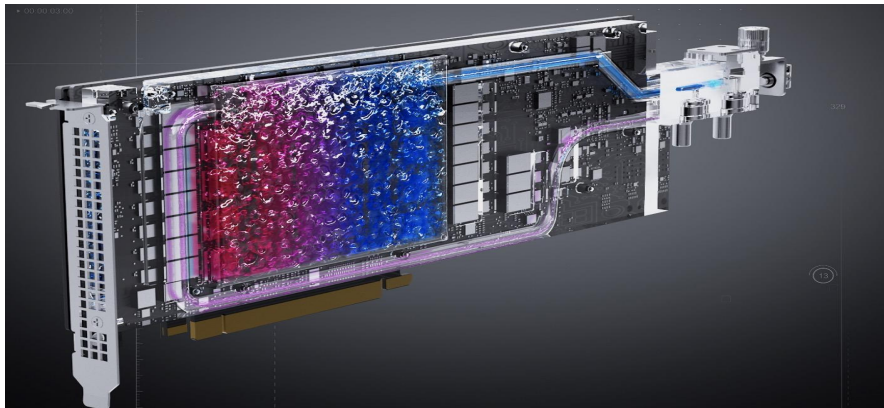
405B-超云OpenLoop冷板液冷八卡机



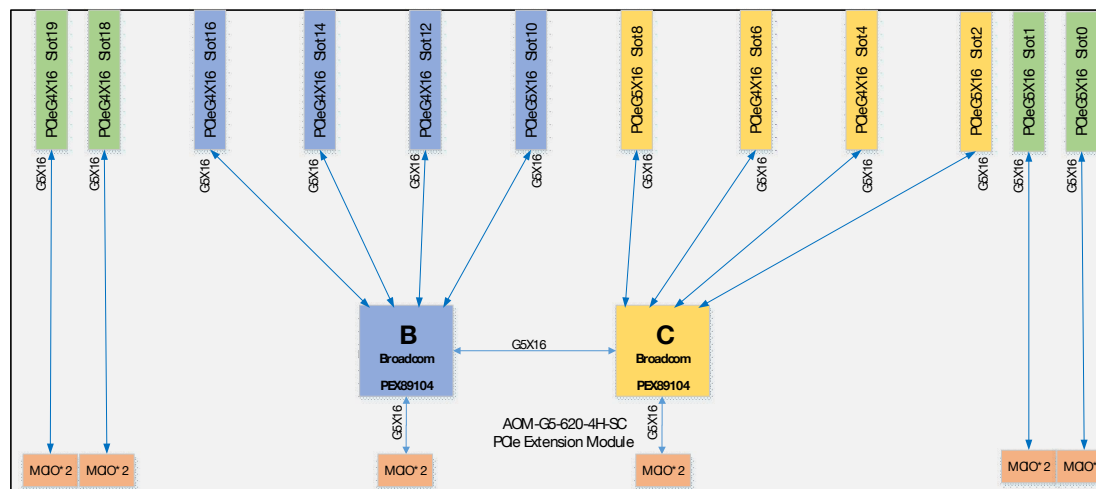
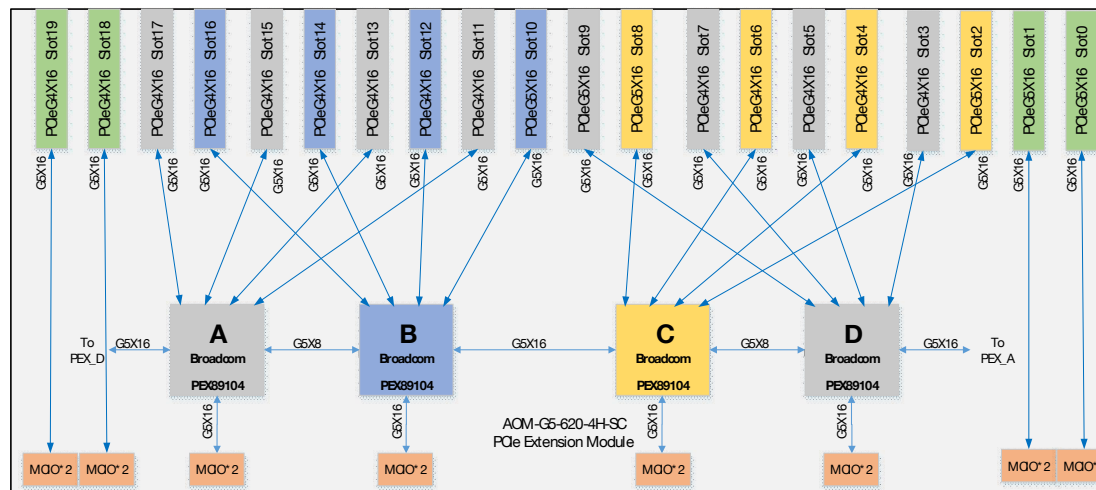
- ❑ 所有槽位支持标卡，GPU槽位最大可支持到非标高157.5以内
- ❑ NV卡与国产GPU均可支持

- ❑ 直通与SW双版本
- ❑ MQD快接应用设计
- ❑ 单卡TDP 支持通流量2+l/min提升至5l/min

671B-超云OpenLoop冷板液冷十六卡机

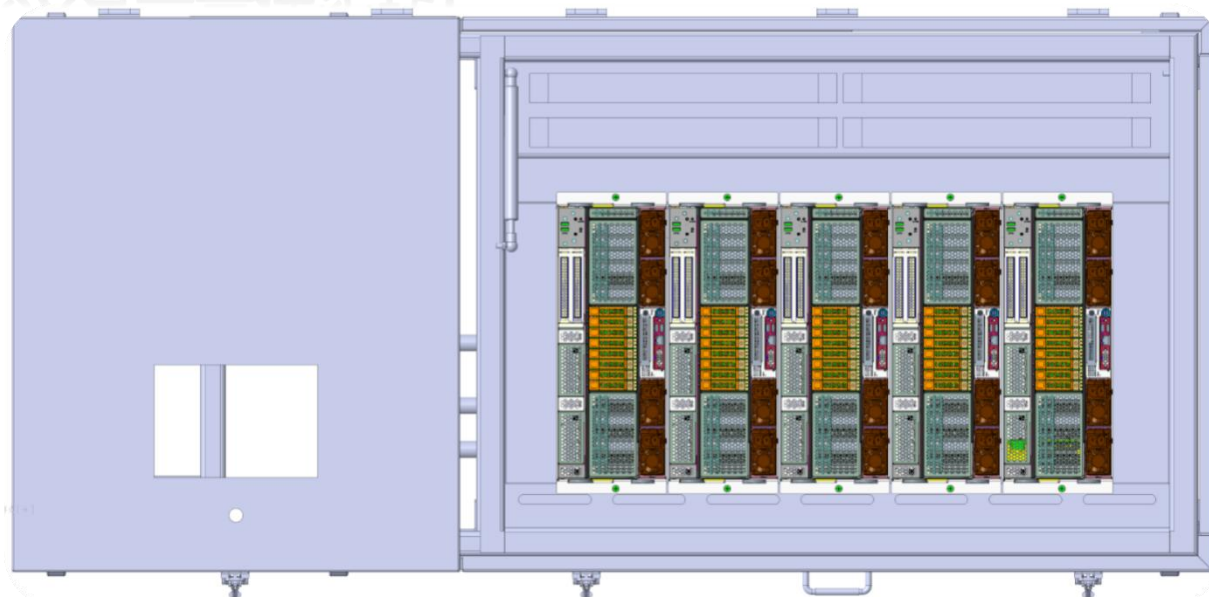


- 多互联通路：16卡SW，NV与国产GPU均兼容
- 高扩展性：20*PCIe5.0 X16



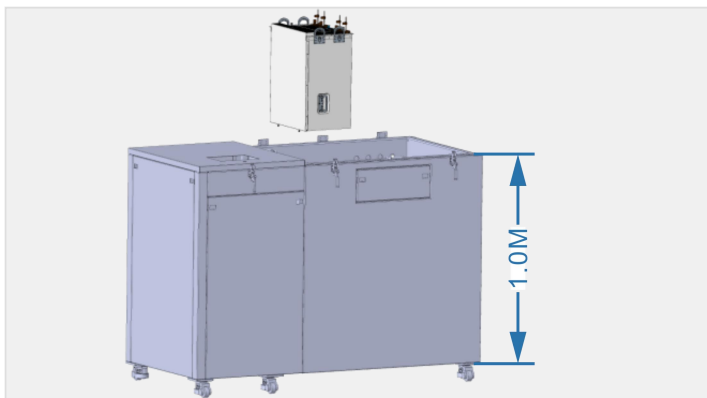
最佳实践-乌兰察布浸没式液冷

携手客户共探浸没式液冷公有云最佳实践!



- 全浸没液冷散热
- 全部件支持前I/O维护
- 处理器支持2颗Intel Xeon® /Emerald Rapids系列处理器，TDP最高350W
- 32根内存插槽，最大可扩展8TB内存；每颗CPU支持8个内存通道；
- 支持高级内存纠错、内存镜像、内存热备等高级功能
- GPU支持8张双宽浸没AI加速卡
- 硬盘支持8个标准2.5英寸热插拔硬盘【预留24SFF盘位】
- 超云4U GPU服务器机箱（长×宽×高）：865×448×175.5（mm）

- 全浸没液冷，纯被动散热，单机成本及功耗5.8kW
- 年运行PUE≤1.05
- 4U支持8卡509X，提供浸没改卡服务
- 前I/O维护，易运维



便捷部署

1. TANK深度1.0m，整机深度565mm
2. 人工可快速提起设备放入TANK内



调试接口

所有调整接口及状态指示灯都在前面板处，TANK出线设计需要兼顾数量、便利性、减少空间占用



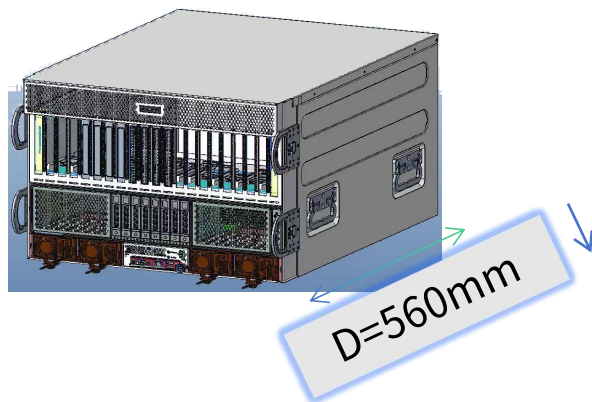
供电接口

1. 极致的高功率密度与散热支持
2. 便利的运维与可维护性

最佳实践-Token出海泰国浸没式算立方

应用场景：

- ✓ 边缘算力中心
空间有限、快速扩容场景
- ✓ 集装箱式数据中心
野外、临时、模块化部署
- ✓ 工业互联网 / 工厂算力站
产线旁 AI 推理、低噪、省空间



项目	产品规格
整机系统（长*宽*高）	3630x885x1088（含脚轮）
服务器1U高度	46.5±0.5mm
水槽可支持	48U
配置支持	超云6U，GPU浸没机型
供电方式	PDU
散热功耗	70kW
系统噪音	55dB
液体	合成油/氟化液
服务器固定	滑轨

浸没式液冷与风冷数据对比

风冷液冷服务器参数对比		
	风冷	浸没
尺寸	6U	4U
重量	65kg	47kg
电源	6pcs	4pcs
风扇	12	NA
导风罩	1pcs	NA
风扇板	2pcs	NA
风扇模组	1pcs	NA

运行数据对比			
对比项	风冷	浸没	差值
功耗	7800w	5800w	2000w
GPU性能	液冷高负载性能更稳定		
CPU性能	基本一致		
GPU温度	75度	63度	12度
散热温度	进风25度	进液35度	
全年电费	68328元	50808元	17520元
液体价格	0	5500	5台机器1年节省下的电费可以买2张509X

液冷与风冷节能对比

	服务器信息	风冷	冷板液冷	浸没液冷
IT设备	单机功耗	7700	7000	5800
	数量	130	130	130
	总功耗kW	1001	910	754

注：浸没液冷服务器本体无动力机械部件，对比8风冷卡机整机功耗可减少约1.9KW。

	基础设施散热部分投资差	风冷PUE	冷板PUE	浸没液冷
基础设施部分	PUE	1.4	1.2	1.05
	基础设施总功耗（千瓦）	1401.4	1092	914.55
	全年电量（度）	12276264	9565920	8011458
	电费（元）	0.87	0.87	0.87
	年运行电费（万元）	1068	832	697
	液冷节电/年（万元）			-236

优势	指标	风冷	冷板式液冷	浸没式液冷
节能	PUE	1.4	1.2以下	1.05以下
	数据中心总能耗单节点均摊	1	0.86	0.75
成本	数据中心总成本单节点均摊	1	0.96	0.74
占地	功率密度（KW/机柜）	10	40	200
	主机房占地面积比例	1	1/4	1/20
CPU可靠	核温（℃）	85	65	65
机房环境	温湿度、洁净度、腐蚀性气体（硫化物、盐雾等）	要求高	要求高	要求低



让AI更简单

谢谢