宝德信创AI产品主打胶片

信创BU

宝德计算机系统股份有限公司

宝德计算机发展历程

宝德 PowerLeader

五大BU齐头并进 & 协同作战

IA BU

- ✓ 通用服务器
- ✓ 大容量存储服务器
- ✓ 高密度四子星服务器
- ✓ AI训练和AI推理服务器
- ✓ 边缘计算产品
- ✓ 图形工作站
- ✓ 存储产品
- ✓ 液冷产品和方案
- ✓ 产品解决方案和行业解决方案

+

信创BU

- ✓ 自强®鲲鹏系列服务器
- ✓ 自强®AI系列训练服务器、中心推理服务器、智能边缘服务器、边缘小站
- ✓ 自强®系列存储
- ✓ 自强®系列台式机、笔记本等终端产品
- ✓ 自强®产品解决方案和行业解决方案

+

网安BU

- ✓ 基于飞腾、龙芯、申威等处理器开发的整机系统
- ✓ 基于Intel处理器开发的整机系统
- ✓ 独特的加固和安全产品
- ✓ 定制化产品

+

宝通BU

- ✓ 为数据中心/云计算/大数据等应用提供硬件设备支持
- ✓ 丰富的物联网产品解决方案
- ✓ 为消费级业务提供产品及服务支持

+

海外BU

- ✓ X86产品方案
- ✓ 自强®产品方案
- ✓ 智算模块&元宇宙等产品方案
- ✓ 为全球用户提供专业优质的产品全生命周期服务

宝德 PowerLeader

宝德国产化业务发展历程

2009

- 自主研发宝德红星OS操作系统；
- 与龙芯、中标正式签署战略合作协议

2010

被工信部授予”优秀安全可控产品云计算解决方案”称号

2011

宝德龙芯单路、双路和存储服务器正式亮相深圳高交会

2012

宝德龙芯双路服务器通过深圳市科技成果鉴定

2013

宝德与深圳大学计算机学院展开合作，开发基于龙芯的高性能四路服务器

2014

宝德与飞腾、麒麟等上下游厂商正式签署战略合作协议，补全整条产业链

2015

宝德推出基于国产飞腾CPU的分布式存储系统

2016

宝德持续加大国产化投入

2017

宝德与飞腾、龙芯联合申报2017年“核高基”项目

2018

宝德申威服务器正式发布

2019

- 宝德自强鲲鹏920系列产品推出；
- 宝德自强飞腾FT2000系列产品推出；
- 宝德自强龙芯系列产品推出

2020

- 南宁、西安鲲鹏产线陆续下线；
- 鲲鹏服务器商业成功；
- 鲲鹏服务器和信创市场的出货量两项NO.1

2021

- 宝德自强AI系列产品推出；
- 宝德自强品牌发布；
- 自强应用生态产业联盟成立；
- 稳居信创整机市场NO.1；
- 自强系列鲲鹏服务器和AI服务器出货量双冠军

2022

宝德自强系列服务器和存储产品布局进一步完善

2023

- 参与天池计划等项目联合开发；
- 自强Mini PC荣获德国红点大奖

2024

运营商、能源、交通、政务等关基行业实现重大突破；

宝德 PowerLeader

研发能力 - 研发创新是宝德发展的第一驱动力

宝德计算拥有服务器主板、部件、BIOS及相关软件系统以及整机系统的自主研发和自主加工能力

荣誉

- 拥有超过600+技术人员
- 专利、知识产权等近200项
- 国家级专精特新“小巨人”企业
- 广东省战略新兴产业骨干企业
- 多次刷新SPEC世界纪录
- 深圳市级企业技术中心
- 广东省级自主安全计算机工程技术研究中心
- 广东省级工业设计中心
- 英特尔-宝德联合培训中心和联合实验室
- 博士后创新实战基地
- 宝德-微软联合体验中心

先进研发设备

雄厚的研发实力

宝德信创业务 - 国之重器，强者自强

自强®鲲鹏务器

优选级鲲鹏整机伙伴

自强®昇腾AI产品

优选级昇腾部件伙伴

华为昇腾万里伙伴计划

大模型一体机伙伴

华为技术有限公司

钻石经销商

NO.20250206CN0478

HUAWEI

认证证书

兹授权：

宝德计算机系统股份有限公司 为

华为技术有限公司

钻石经销商

授权直接下单产品：计算

授权经销产品：企业服务与软件;光传送;光接入;华为云 Stack;协作平板;存储;数通;数通网络安全;智能安防;视频会议;计算

授权经销区域：中国大陆地区（不含港、澳、台）

授权经销行业：所有行业（不含运营商及已有排他性授权的行业）

授权经销模式：仅限线下授权，不得转授权

授权经销期限：2025 年 02 月 02 日至 2026 年 02 月 01 日

华为技术有限公司

昇腾持续打造极致性能、极简易用的全场景人工智能平台

行业应用	运营商、互联网、能源、金融、交通、制造、医疗等行业应用

应用使能	ModelArts	HiAI Service		第三方平台
	MindX 昇腾应用使能			
	MindX DL	MindX Edge	ModelZoo	MindX SDK

全流程开发工具链

Mindstudio

AI框架	[M]S 昇思 MindSpore	TensorFlow/PyTorch 等第三方框架
异构计算架构	CANN	

Atlas系列硬件

推理系列


Atlas 200I A2
AI加速模块


PI300T G2
智能小站


Atlas 300V


Atlas 300I/V Pro


Atlas 300I Duo
推理卡


PR205KI智能
边缘服务器


PR210KI系列
推理服务器

训练系列


Atlas 300T Pro
训练卡


PR420KI G2
系列训练服务器


PRA100 AI 集群


PRA100 PoD G2

昇腾加速卡介绍

	边缘推理			视频推理		中心推理				中心训练		
	Atlas 200I A2 (8T)	Atlas 200I A2 (20T)	Atlas 200I DK A2开发者套件	Atlas 300V	Atlas 300V Pro	Atlas 300I	Atlas 300I Pro	Atlas 300I Duo	Atlas 300I A2	910B_4	910B_3/2	910C
FP16 (TFLOPs)	4	10	4	50	70	44	70	140	280	280	313/376	560
INT8 (TOPs)	8	20	8	100	140	88	140	280	560	560	626/752	1120
显存容量	4GB LPDDR4X	4/8/12GB LPDDR4X	4GB	24GB LPDDR4X	48GB LPDDR4X	32GB LPDDR4X	24GB LPDDR4X	96GB LPDDR4X	32/64GB HBM	32/64GB HBM2e	64GB HBM2e	128GB HBM2e
显存带宽	25.6GB/s	34.1/51.2GB/s	25.6GB/s	200GB/s	300GB/s	200GB/s	200GB/s	400GB/s	800/1600 GB/s	800/1600 GB/s	1600GB/s	3200GB/s
视频解码/路 1080P 30FPS	20	40	16	80	128	54	100	256	480	/	/	/
图片解码/张	512 (1080P)	512 (1080P)	512 (1080P)	384 (4k)	384 (4k)	256 (4k)	384 (4k)	768 (4k)	/	/	/	/
互联方式	高速SerDes	高速SerDes	USB	PCIe 4.0 x16	PCIe 4.0 x16	PCIe 3.0 x16	PCIe 4.0 x16	PCIe 4.0 x16	PCIe 4.0 x16	392GB/s HCCS	392GB/s HCCS	784GB/s HCCS
功耗	21W	25W	24W	72W	72W	67W	72W	150W	300/350W	350W	365/380W	/
形态	模块	模块	套件	半高半长单宽	半高半长单宽	半高半长单宽	半高半长单宽	全高全长单宽	全高全长双宽	HAM模组	HAM模组	HAM模组
散热方式	被动散热	被动散热	被动散热	被动散热	被动散热	被动散热	被动散热	被动散热	被动散热	被动散热	被动散热/液冷	被动散热

华为昇腾芯片演进路线

NPU芯片型号	Acend 910B	Ascend 910C	Ascend 950PR	Ascend 950DT	Ascend 960	Ascend 970
发布时间	2023Q1	2025 Q1	2026Q1	2026Q4	2027Q4	2028 Q4
指令架构	SIMD	SIMD	SIMD/SIMT		SIMD/SIMT	SIMD/SIMT
支持精度	FP32/FP16/BF16/INT8	FP32/HF32/FP16/BF16/INT8	FP32/HF32/FP16/BF16/FP8/MXFP8/HiF8/MXFP4/		FP32/HF32/FP16/BF16FP8/MXFP8/HiF8/MXFP4/HiF4	FP32/HF32/FP16/BF16FP8/MXFP8/HiF8/MXFP4/HiF4
HCCS互联带宽	392 GB/s	784 GB/s	2 TB/s		2.2 TB/s	4 TB/s
AI性能	400/376/313/280 TFlops FP16	800TFlops FP16	1 PFlops FP8，2 PFlops FP4		2 PFlops FP8, 4 PFLops FP4	4 PFlops fp8, 8 PFlops FP4
显存容量带宽	64 GB, 1.6TB/s	128 GB, 3.2TB/s	128 GB, 1.6TB/s	144 GB, 4TB/s	288 GB, 9.6TB/s	288 GB, 14.4TB/s

昇腾产品2025整体规划，打造云边端训推解决方案

宝德昇腾人工智能全产品线

昇腾训练服务器

A+K
液冷智能
计算集群

PRA100 PoD G1/G2/G3

A+K
训练服务器

PR420KI G1/G2/G3

昇腾推理服务器

A+X
推理服务器

PR2715E

PR4908E

PR410EI

A+K
推理服务器

PR210KI /PRO

PR410KI

PR425KI G2

昇腾边缘推理

A+X
推理工作站

PT6610A

A+K
边缘推理
服务器

PR205KI

昇腾边缘模组

智能小站

PI300T

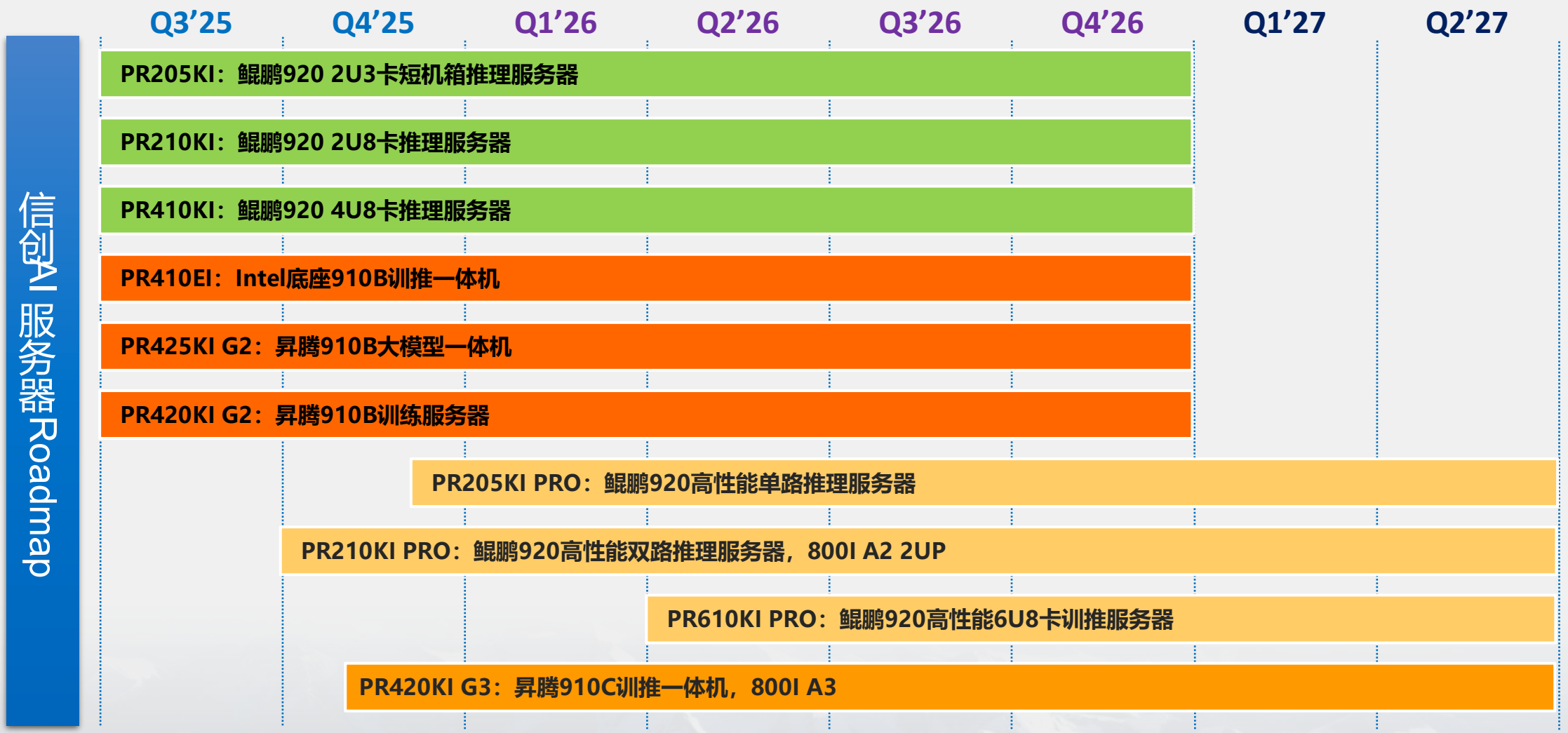
PI300T G2

Atals 200I DK

Atals 200I A2

宝德 PowerLeader

宝德信创AI服务器Roadmap



宝德大模型训推一体机 PR420KI G3

关键特性	规格描述
形态	10U机架服务器（442mm(H) × 447mm(W) × 790mm(D)）
CPU	4 X 鲲鹏920处理器，7280Z 80C/2.9G，7285Z 80C/3.0G
内存	支持32个DDR内存插槽，最大支持5200MHz，支持32G/64GB
NPU	8 X 昇腾910C NPU，单NPU显存**128GB**
AI算力	单NPU AI算力 **560** TFlops（FP16），总AI算力 **4.48** PFlops
内部拓扑	NPU HCCS全互联，互联带宽782GB/s
网络接口	8 * 400GE QSFP接口直出，RoCE网络协议（ROCE方案） 56 * 400GE QSFP接口直出，灵衢总线协议（灵衢方案）
功耗散热	最大功耗14.6KW，风冷散热，净重229kg

上市信息 已上市　　**当前状态** **可销售**

应用场景

面向大模型一体机场景产品升级，更大算力，更大内存：全面适配大模型微调、集群推理；

互联网

科研教育

行业大模型

AI算力

560T VS 148T

3.78x H20

大模型推理超越友商顶级卡

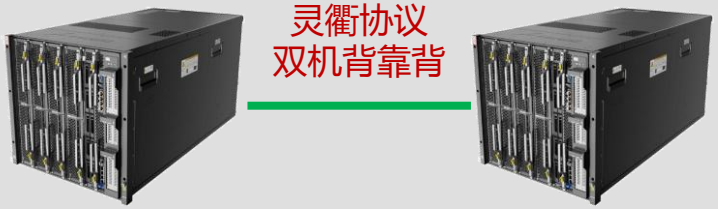
宝德 PowerLeader

宝德PR420KI G3超节点组网方案

单机超节点标准方案

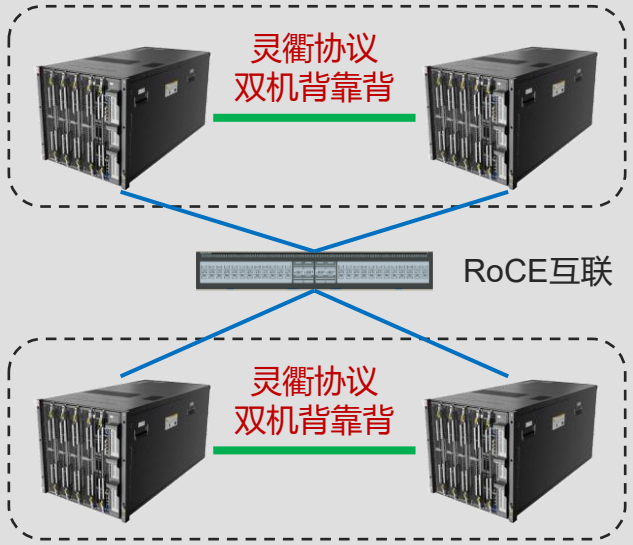
1. 一台为单位销售，RoCE方案；
2. 支持RoCE互联，不支持跨服务器超节点互联；

双机超节点直连方案



1. 2台为单位销售，灵衢方案；
2. 支持2台之间用灵衢线缆直连；

多机超节点组网方案



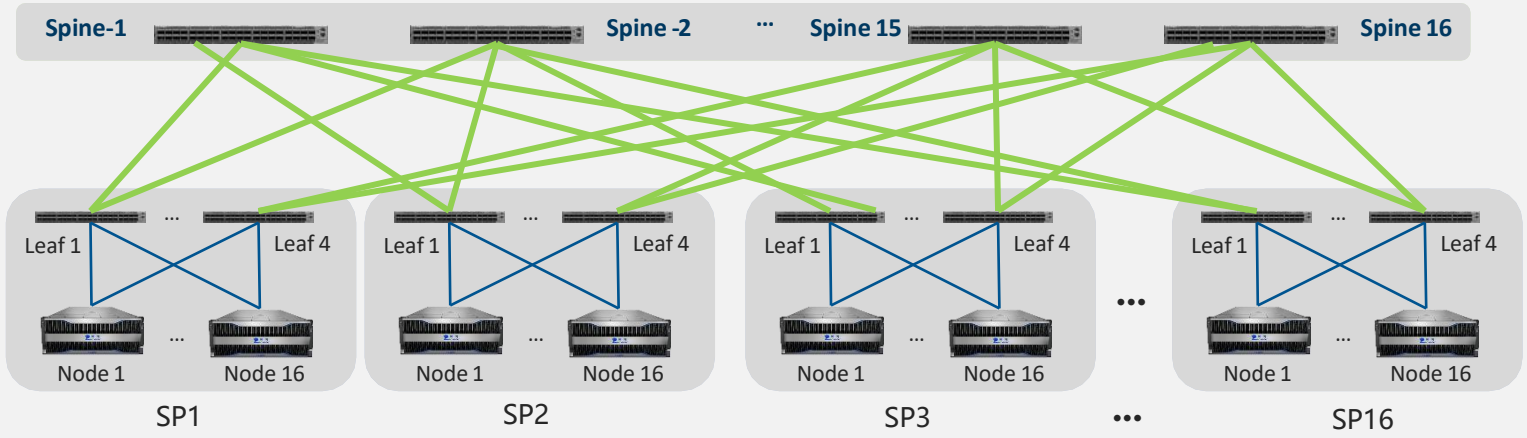
1. 双机背靠背+多组双机RoCE互联；
2. 多台灵衢方案+RoCE交换机组网；

宝德昇腾中心训练产品（PR420KI G2）

关键特性	规格描述	
形态	4U机架服务器（175mm × 447mm × 790mm）	
CPU	4 * 鲲鹏920	
NPU	8 * 昇腾910	
AI算力	半精度（FP16）**2.5/3.0** PFLOPS	单精度（FP32）**0.65/0.8**PFLOPS
内存	8 * 64G HBM；支持32个DDR4内存插槽	
内部拓扑	NPU HCCS全互联，互联带宽392GB/s	
网络接口	NPU直出8 * 200G RoCE	

上市信息 已上市　　**当前状态** **可销售**

Pod 集群组网



系统参数	性能
峰值AI算力	641PFlops FP16
跨结点带宽	跨POD任意两节点互联带 宽200GB/s
扩展性	16个POD，2048颗NPU

以256台训练服务器为例，共计2048颗Ascned 910芯片
分成16组，每组16台服务器，里面放置一个完整的模型，每组之间采用数据并行；每组16台服务器，采用模型并行，每台服务器放1/16的模型；

宝德A+X训推一体机PR410EI

CPU	2 x Intel Sapphire Rapids CPU
AI算力（910B4）	（FP16）**2.24P**FLOPS
HBM	32G 800GB/s / 64G 1.6TB/S
节点内互联	HCCS 392GB/s
跨节点互联	8*200G Roce

CPU	2 x Intel Sapphire Rapids CPU
AI算力（910B3/2）	（FP16）**2.5~3.0P** FLOPS
HBM	64G 1.6TB/S
节点内互联	HCCS 392GB/s
跨节点互联	8*200G Roce

上市信息 已上市　　**当前状态** **可销售**

应用场景　面向互联网/大模型/科研教育等市场，全面适配大模型微调、集群推理；

互联网

科研教育

行业大模型

直通方案:

- PICE插槽：最大支持5个PCIe5.0x8；
- 存储：6*NVME+4SATA/SAS 或 4*NVME+6SATA/SAS 或 2*NVME+8SATA/SAS

Switch方案:

- PR410EI支持4个博通PCIe5.0 Switch，每个PCIe Switch通过2组PCIe 5.0 x16与CPU对接；
- 最大支持10*NVME，支持4个PCIe 5.0 x16、2个PCIe5.0 x8扩展插槽；

宝德 PowerLeader

宝德昇腾DS满血版推理产品（PR425KI G2）

上市信息 已上市　　**当前状态** 可销售

关键特性	规格描述	
形态	4U机架服务器（175mm × 447mm × 790mm）	
CPU	4 X 鲲鹏920	
NPU	8 X 昇腾NPU	
AI算力	FP16半精度 **2.24** PFLOPS	FP32单精度 **0.60** PFLOPS
显存	**8 X 64G 1600Gb / 8 X 32GB 800Gb HBM**；支持32个DDR4内存插槽	
内部拓扑	NPU HCCS全互联，互联带宽392GB/s	
网络接口	NPU直出 8 X 200G RoCE	
散热	风冷散热	

应用场景

面向互联网/大模型/科研教育等市场，全面适配大模型微调、集群推理；

互联网

科研教育

行业大模型

910B4_64G 主要面向Deepseek-R1 671B满血版集群推理

910B4_32G 主要面向Deepseek 70B以下蒸馏版模型推理

宝德 PowerLeader

宝德昇腾DS蒸馏版推理产品（PR410KI）

上市信息 已上市　　**当前状态** 可销售

应用场景 部署于数据中心机房中，使能AI中心推理

精准营销	视频分析	OCR	医疗影像分析
智慧零售	智慧城市	智慧金融	智慧医疗

关键特性	规格描述
形态	4U机架服务器（175mm × 448mm × 800mm）
CPU	2 X 鲲鹏920，支持32/48/64核主板
内存	32 X DDR4内存插槽，支持16GB/32GB/64GB
硬盘	前置4 X 2.5英寸SATA/SAS+8 X 2.5英寸NVMe；
AI加速卡	最大支持10张全高全长单/双宽标卡； 最大支持10 张Atlas 300I Duo， **1.4P** AI算力，960GB显存； 最大可支持8*Atlas 300I A2卡，单卡AI算力280TFlops，32GB 800GB/s或64GB 1600GB/s显存； 最大支持8张MX C500推理卡，512GB HBM显存；
PCIe扩展	最多支持13个PCIe 4.0 标准扩展插槽
电源	4个热插拔2000W电源模块，支持2+2/3+1冗余

面向人工智能市场，CPU/GPU全栈自主创新，AI 性能领先

4U多卡服务器方案，面向DS 32B/70B场景单机支持更多并发用户

宝德 PowerLeader

宝德昇腾DS蒸馏版推理产品（PR210KI）

Atlas 300I
Atlas 300I Pro
Atlas 300V Pro

支持 **1 ~ 8** 张卡

Atlas 300I Duo 推理卡

支持 **1 ~ 4** 张卡

插卡式

自强 PR210KI 推理服务器

Atlas 300I
Atlas 300I Pro
Atlas 300V
Atlas 300V Pro

支持 **1 ~ 7** 张卡

插卡式

自强 PR215PI 推理服务器

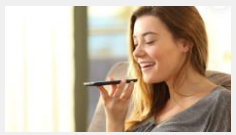
上市信息 已上市　　**当前状态** **可销售**

应用场景 部署于数据中心机房中，使能AI中心推理

搜索推荐

金融大脑

语音识别



内容审核

关键特性	规格描述	
型号	PR210KI	PR215PI
形态	2U机架服务器 86.1mm × 447mm × 790mm	2U机架服务器 86.1mm × 447mm × 748/708mm
CPU	2 x 鲲鹏920	2个Intel® Xeon® Purley处理器
内存	最多32个DDR4内存插槽	最多24个DDR4内存插槽
AI加速卡	最大支持 **8** 张Atlas 300I/V、300I Pro/300V Pro， 或 **4** 张Atlas 300I Duo 96GB	最大支持 **7** 个Atlas 300I、300IPro/300V Pro

Atlas 300I Duo卡面向DeepSeek 32B及以下蒸馏模型推理

宝德 PowerLeader

Atlas 300I A2 推理方案：极致性能，使能大模型推理

关键特性	规格描述
形态	双槽位全高全长PCIe卡
尺寸	266.7mm（长）× 39.04mm（宽）× 111.15mm（高）
处理器核	内置 20个AI Core ， 8个Arm Core
显存	容量：显存**32GB 800GB/s**、 **64GB 1600GB/s HBM2**，支持ECC
AI算力	整数精度（INT8）：**560** TOPS 半精度（FP16）：**280** TFLOPS
编解码能力	内置DVPP预处理单元，支持1080P 480 FPS等效的视频解码能力（硬件解码能力）
功耗	300/350 W

上市信息 已上市

应用场景

- 集成于服务器，进行AI推理
- 大模型推理场景、搜索推荐等场景

优势

- 单卡性能高，对标业界厂商时，在互联网常用搜索NLP、推荐、内容审核等场景性能优势显著
- 大模型推理，对标友商顶级卡，性能比肩，性价比持平

大模型推理

AIGC（对话、语音、以文生图...）

互联网搜索推荐

推荐、搜索NLP、内容审核

1 大模型推理比肩友商顶级卡

Atlas 300I A2 vs 业界厂商

性能 **0.8~1.0X** 性价比持平业界厂商

2 互联网场景性能强 搜索NLP、推荐、内容审核

Atlas 300I A2 vs 业界厂商 **1.5~2X↑**

宝德 PowerLeader

宝德A+X中心推理服务器（PR4908E/PR2715E）

Atlas 300I DUO 推理卡

支持 **1 ~ 8** 张卡

插卡式

自强PR4908E推理服务器

Atlas 300I Pro
Atlas 300V Pro

支持 **1 ~ 6** 张卡

Atlas 300I DUO 推理卡

支持 **1 ~ 4** 张卡

插卡式

自强PR2715E推理服务器

关键特性	规格描述	
型号	PR4908E	PR2715E
形态	4U机架服务器	2U机架服务器
CPU	2 x Intel第4/5代至强CPU	2 x Intel第4/5代至强CPU
内存	32 x DDR5 内存插槽	32 x DDR5 内存插槽
AI加速卡	最大支持 **8** 张Atlas 300I DUO **1.12P** AI算力，**768GB**显存	最大支持 **6** 个Atlas 300I/300IPro/300V Pro，或 **4** 张Atlas 300I DUO
散热	风冷散热	

典配策略	配置参考
单机推理典配2	PR4908E：2*6530 (32C/2.1GHz) + 16*64GB +2*960GB/2*3.84T U.2/1*9560-8i + 8*Atlas 300I Duo + 2*25G双口
单机推理典配3	PR2715E：2*5418Y (24C/2.0GHz) + 16*32GB +2*960GB/4*8TB SATA/1*9560-8i + 6*Atlas 300I Pro + 2*25G双口

宝德 PowerLeader

宝德昇腾边缘推理产品PR205KI：满足“短机柜”部署

Atlas 300I

Atlas 300I Pro

Atlas 300V

Atlas 300V Pro

支持 **1~3** 张卡

插卡式

*无Riser卡的情况

“短机柜”

室外电信设备机柜（III、IV）

典型尺寸深度 **475mm**

PR205KI 智能边缘服务器

关键特性	规格描述
形态	2U服务器，短机箱（86.1mm x 447mm x **475mm**）
AI加速卡	最大支持 **3张** Atlas 300I/V Pro 推理卡
CPU	1 x 鲲鹏920 （单路24核）
内存	4个DDR4内存插槽，最高3200 MT/s
AI算力	最大 **420** TOPS INT8 或 **384** 路1080P 30FPS视频解析（硬件解码能力）

应用场景

边缘侧独立部署，使能智能边缘

大型园区、电力、交通、商超等场景

智慧交通

湖南高速

变电站智能巡视

国网、南网

智慧加油站

加油站生产监测

1 稳定算力高

PR205KI VS 主流产品 **1.6X** ↑

TOPS (INT8)

2 能效比高

PR205KI VS 主流产品 **2.6X** ↑

TOPS/W (INT8)

3 视频解码能力强

PR205KI VS 主流产品 解码路数 **2.4X** ↑

4 典型模型性能好

1.9X ↑ **业界友商** images/s

1.3X ↑ **业界友商** sentences/s

宝德 PowerLeader

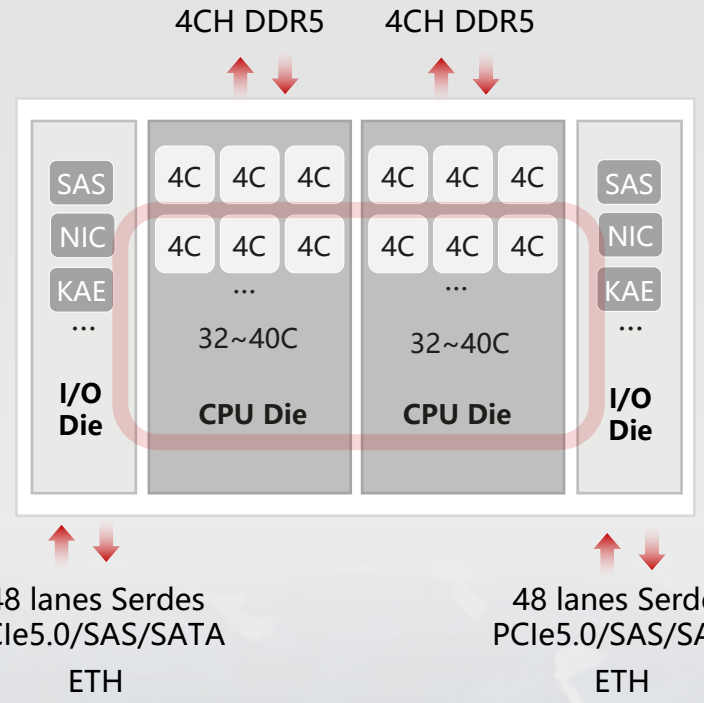
鲲鹏920新型号处理器，完全自主设计，性能、规格大幅度提升

核心技术完全自主

自研处理器内核

自研片内互联协议

自研片间互联协议



VS. 鲲鹏920

- SPECINT2017性能：**2.6倍**
- 单核性能：**1.6倍**
- 整芯片FP64浮点性能：**5.6倍**
- 内存带宽：**1.63倍**（DDR4 -> DDR5）

鲲鹏920新型号规格

- 支持**64/80核等**多种规格，主频支持2.9GHz等频率
- 支持**双线程SMT2**
- 8个**DDR5通道**，支持**速率4800MT/s及以上**
- **集成96 Lane PCIe 5.0**
- **L2 cache 1.25MB/core ， L3 cache 1.75MB/core**
- 支持**2*200G** 、4*25/100GE、RoCEv2、标准NIC网卡
- 支持2P/4P互联
- 集成硬加密和压缩引擎

** 基于鲲鹏7280Z vs鲲鹏7260处理器的对比数据*

宝德 PowerLeader

鲲鹏920新型号8大升级，构建高性能、高能效、高可靠、高安全4大能力

高性能：微架构、高速I/O、浮点、KAE加速全面升级

1 微架构升级

L2 cache/L3 cache提升，流水线优化

单核性能是920的**1.6倍**

首个**双线程**ARM处理器

2 高速I/O升级

PCIe4.0 —> PCIe5.0

DDR4 —> DDR5

3 浮点能力升级

支持SVE，**独有2*256位宽**

FP64浮点算力是920的**5.6倍**

4 KAE加速能力升级

国密支持SM2、SM3、SM4

加速引擎：ZLIB、GZIP、ZSTD等

高能效：能效设计全面增强

5 能效升级：从处理器、主板到整机全栈联动节能

处理器

动态调频调压，节能最高10%+

主板

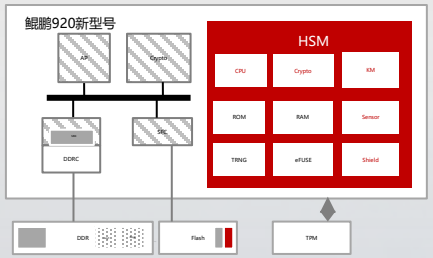
精准调压，板级功耗降低~2%

整机

极致散热设计，整机功耗降低~7%

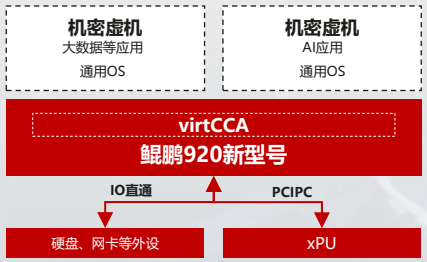
高安全：安全启动增强，独有支持异构安全

6 支持高安启动，独有硬件三防



硬件三防：防侧信道、防故障注入、防物理攻击等

7 virtCCA+PCIPC 设备直通安全域



高可靠：RAS能力升级，独有支持内存镜像能力

RAS1.0

CPU核故障在线隔离

内存故障自愈

内存单bit纠错

…

8 RAS2.0

内存镜像

CPU核故障在线隔离

内存故障自愈

内存单bit纠错

…

宝德 PowerLeader

昇腾最新一代推理服务器PR210KI PRO，CPU全新升级

支持 **1 ~ 8** 张卡

Atlas 300I
Atlas 300I Pro
Atlas 300V Pro

插卡式

支持 **1 ~ 6** 张卡

Atlas 300I Duo 推理卡

自强 PR210KI PRO 推理服务器

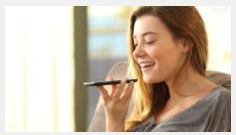
上市信息 已上市　　**当前状态** **可销售**

应用场景 部署于数据中心机房中，使能AI中心推理，支持大模型RAG

搜索推荐

金融大脑

语音识别



生成式内容

关键特性	规格描述
形态	2U机架服务器（86.1mm × 447mm × 798.5mm）
CPU	2 X 鲲鹏处理器，支持7270Z 64C2.9GHz、7280Z 80C2.9GHz
内存	32 X DDR5内存插槽，支持32GB/64GB 5200MHz/4800MHz；
硬盘	前置8 X 2.5英寸SATA/SAS，兼容4 X 2.5英寸NVMe；
AI加速卡	最大支持4张Atlas 300I A2，1.12P FP16 AI算力，128GB HBM显存； 最大支持6张Atlas 300I Duo，**1.68P** INT8 AI算力，476GB显存；
PCIe扩展	最多支持8个PCIe 5.0 标准扩展插槽
电源	2个热插拔2000W/2600W电源模块，支持1+1冗余

AI算力密度

给客户提供高性能、高性价比的GCH推理标卡底座，提供稳定可靠的推理服务器产品

宝德 PowerLeader

PI300T G2：边缘宽温部署，视频智能分析利器

3.5英寸硬盘仓

4T / 8T / 16T

PI300T G2 智能小站

上市信息 已上市　　**当前状态** 可销售

PI300T 演进一代产品

应用场景 满足严苛的边缘部署场景设计，在智慧城市、交通、社区、园区、商场、超市等复杂环境区域应用

智慧网点	自由流收费	智慧园区	智慧加油站
招行、工行	全国ETC收费站	商汤	江苏石化



关键特性	规格描述
尺寸（长x宽x高）	无盘配置：290 mm x 220 mm x 44 mm 有盘配置：410 mm x 220 mm x 44 mm
内存	LPDDR4X，12GB / 4GB； 总带宽51.2 GB/s
AI算力	整数精度（INT8）：**20**TOPS　半精度（FP16）：**10**TFLOPS
编解码能力	内置DVPP预处理单元　图片JPEG/PNG编解码能力 视频 **40路** 1080P 30FPS
重要接口	2*USB3.0；**5***RJ45千兆网口；1* M.2 KEYB（可接**5G**模组）； 4*DI / 4*DO；（选配：1*MicroSD卡槽；1*M.2 NVMe SSD）
媒体	2路HDMI图片输出，满足现场结果直接显示
典型功耗	无盘 32.3W / 有盘 44.5 W
环境条件	-40℃ ~ +60℃

优势

- 重点引导视频分析能力强、接口丰富适应更多边缘场景

1 视频分析能力强 1080P 30FPS

PI300T G2	VS	友商
最大40路		最大32路

1.25X

2 接口更丰富

PI300T G2	VS	友商
5路	千兆网口	1路
2路	HDMI	1路

宝德 PowerLeader

Atlas 200I DK A2：开箱即用、参考丰富的开发者套件

Atlas 200I DK A2
开发者套件

上市信息 已上市　　**当前状态** 可销售

Atlas 200 DK
演进一代产品

早期小批量发货

关键特性	规格描述
形态	135mm x 120mm x 44mm
AI算力	整数精度（INT8）：**8/20** TOPS　半精度（FP16）：**4/10** TFLOPS
内存	容量4GB（8T算力）、8GB/12GB（20T算力）
摄像头接口	2*MIPI-CSI 支持两个树莓派摄像头
USB接口	1*USB TypeC 仅支持从模式；　2*USB3.0 Type-A
以太网接口	2*RJ45千兆网口
编解码能力	内置DVPP预处理单元　图片JPEG/PNG编解码能力 视频 **16路** 1080P 30FPS
功耗	24 W
存储设备	1*NVMe/SATA M.2 SSD；　1*Micro SD卡
图形显示	2* 4K分辨率HDMI视频输出

应用场景 昇腾AI开发者上手学习，实践创新场景，提供配套软硬件

AI应用创新
机械人/智能小车...

高校教学
智能基座

行业算法验证
工业互联网等ISV

1 丰富代码示例

- 3大典型场景示例，覆盖开发者80%应用场景

智能车　机械臂　语音交互

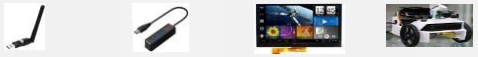
2 开源预训练模型库

- **ModelZoo**：900+高性能预训练模型，CV/NLP/语音等

3 专业认证硬件配件

认证配件，降低选型兼容难度

连接扩展　接口扩展　组件扩展　场景扩展



4 高效、易用开发工具

- **一键制卡工具**：一键安装镜像，30min环境搭建
- **模型适配工具**：Windows下的端侧模型适配工具，支持用户完成端侧模型适配开发全流程

Atlas 200I A2： “以一当四” SoC，超强边缘视频分析

Atlas 200I A2 加速模块

关键特性	规格描述
形态	采用MXM连接器： 82 mm（长） * 60 mm（宽） * 7 mm（高）
内存	20 TOPS： LPDDR4X， 12 GB/8GB/4GB，总带宽 51.2/34.1/34.1 GB/s 8 TOPS： LPDDR4X，4 GB，总带宽 25.6 GB/s
AI算力	整数精度(INT8)：20 /8 TOPS 半精度(FP16)：10 /4 TFLOPS
编解码能力	内置DVPP预处理单元 图片JPEG/PNG编解码能力 视频 40路 1080P 30FPS
典型功耗	20 TOPS： 25 W / 8 TOPS： 21 W
ISP	(新增功能) 4KP45（4K画质每秒45帧） / 8KP15 / 16KP4
图形显示	(新增功能) 2路HDMI图片输出，满足现场结果直接显示

Atlas 200 演进一代产品

应用场景 嵌入边缘设备， 如机工控机、智能网关等室内、桌面级部署的边缘AI场景，大型机器人等供电充足应用场景， 安防场景边缘后端(非太阳能供电)

智能边缘设备
研扬、飞途昇腾等IHV

智能机器人
达闼科技等

视频分析
园区、社区

集成度高 以一当四，不需单配CPU和ISP， 硬件性价比更高

Atlas 200I A2

等于

CPU	NPU	编解码	图形显示
4核LX900 1.6GHz	20/8 TOPS INT8	最大40路1080P@30fps 2路8K@30fps 业界最强	单核Mali-G52 2 路 HDMI

1 视频性能强，单路性价比高

Atlas 200I A2 VS 业界厂商

INT8 算力	硬件视频解析	SoC内存
持平	3x ↑	1.5x ↑

2 接口丰富，行业场景更容易适配

Atlas 200I A2 VS 业界厂商

MXM接口 (314pin) DDR4 SO-DIMM接口 (260pin)

伙伴硬件实现更多的网络/摄像机/USB/串口/音频等

宝德 PowerLeader

宝德昇腾智算中心集群解决方案

应用层

训练推理　远程协作　自然语言　数据分析　机器学习　图像处理

AI计算加速平台

异构计算平台CANN

全流程开发工具链MindStudio　AI框架昇思MindSpore　昇腾推理引擎MindIE

AI资源管理调度平台

PLStack　分布/并行计算优化　异构资源调度与编排　自动化交付

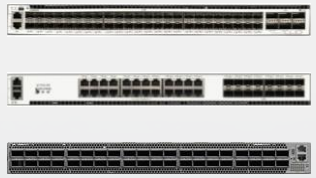
AI基础设施平台

Kunpeng

Ascend NPU

宝德通用服务器

宝德昇腾服务器



宝德交换机

宝德 PowerLeader

宝德智算平台 PLStack 特色功能 – 管理服务

平台管理

- **多租户**管理、多种计量计费模式；
- 集群资源管理：以**套餐方式**进行资源配额；
- 自动化运维工具：集群节点扩缩容；
- 集群监控与告警：安全管理和完善的日志审计功能。

GPU卡管理

- **GPU卡四种模式**：独享/共享/vGPU/MIG分配模式；
- **vGPU**：将GPU卡切分出多张小的vGPU卡，对其显存和算力进行限制，提高物理GPU卡利用率；
- **虚拟显存**：物理显存不足时，通过对显存做虚拟化处理，使得可用显存超过物理显存，从而支持大批量、大模型的训练任务。

多数据中心管理

- 支持将多个物理区域的GPU资源**统一纳管**；
- 统一对多个区域资源使用监控、计量计费等管理；
- 用户可选择不同区域的资源并调用；
- **优化成本**：降低对运维人员成本投入。

数据管理

- **在线标注工具**：使用标注数据进行模型开发、训练、预测；
- **存储管理**：持久化存储工作目录、可视化文件管理系统、共享存储；
- **存储性能、安全性**：分布式文件存储、将本地硬盘组建分布式存储。

宝德智算平台 PLStack 特色功能 – 模型开发训练

开发环境

- **一键式环境生成**，集成数十种集成主流AI框架，如TensorFlow、Pytorch、PaddlePaddle等，支持自定义框架镜像；
- **Mlab**交互式开发工具：兼容Jupyter；
- 定时快照/备份，数据快速回滚；
- 弹性伸缩，资源/框架/存储弹性变更。

模型训练

- 分布式训练：深度集成Horovod、Ray的分布式并行训练，支持**多机多卡，单机多卡**；
- 参数调优：内置Auto ML自动调优，提高模型训练效率；
- 训练可视化，训练过程中，实时输出资源的利用率和模型训练日志。

模型管理

- 集成行业预训练模型和行业数据集，降低用户模型开发难度；
- 多版本模型管理：训练模型和本地模型一键导入；
- **格式转换**：模型文件支持一键转换为ONNX格式；
- 云端服务：模型发布为云端服务能力，并对外提供http协议访问接口。

模型服务

- 模型在线推理应用：直接上传推理文件进行**在线模型推理**；
- **远程调用**：发布后的模型服务提供对外的Web接口和Token秘钥，实现模型的真正应用。
- 数据回传：算法应用后的数据资源可反哺数据存储中，方便后续持续训练更新；

谢谢

国之重器 强者自强