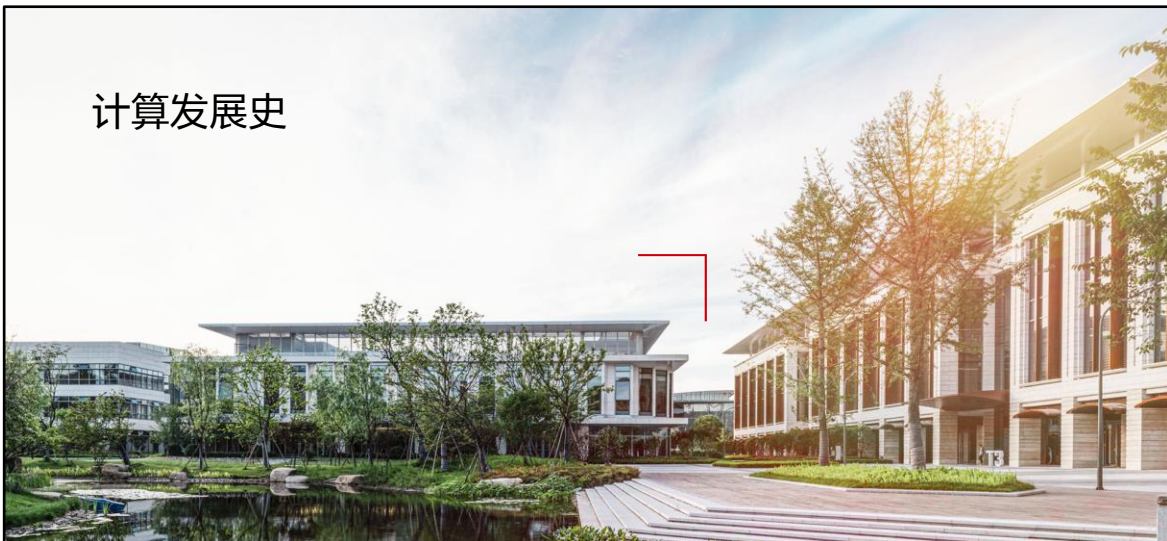


计算发展史



前言

- 随着社会加速步入到智能计算的时代，信息技术领域的推陈出新变得越来越快，以云计算、大数据、人工智能为代表的计算产业也吸引了很多之前不太了解计算产业的人们的目光，为了帮助初学者了解计算是如何从古代一步步走到现在，本章将着重介绍计算的发展历程和发展趋势。

目标

- 学完本课程后，您将能够：
 - 了解计算发展历程
 - 了解计算现在和未来发展趋势

目录

1. 计算在生活中的应用
2. 计算发展历程
3. 计算发展趋势

人们眼中的计算

- 提到计算，人们很容易首先想到数学计算，通过使用各种各样的函数，公式去解决各类数学问题。然后再联想到计算器，人们使用计算器来辅助计算，比如计算生活中的各类收入和支出从而制作账单。在这之后我们还可以想到我们常用的各类电子设备，例如手机和电脑，它们的出现极大的改变了人们工作生活的方式，获取和处理各类数据的方式。



- 不过计算内涵的就仅仅包括这几个方面么？除了这些之外，在人们日常的工作和生活中，我们的身边还有哪些场景能够体现计算为我们的生活带来了便利呢？

我们身边的计算

- 人们现在使用的各种各样的信息技术服务其实也是现代计算的一个典型体现。



- 在线电商，预定酒店，即时通信等等人们最常使用的各类便利的服务都是建立在现代化的计算基础上的，那人类又是如何一步步的从古代开始，将计算的内涵从只是进行简单的数学运算，发展至利用高性能服务器和计算机来构建现在的信息化世界的呢？让我们通过后续的章节来一同了解一下计算的发展历程。

- 身边的IT技术正在改变我们的生活，例如通过APP承接、下发订单的打车软件，通过互联网进行实时语音通话的通讯软件，通过APP提供网上购物体验的电子商城等。
- 出租车行业：UBER、滴滴。
- 酒店行业：Airbnb。
- 电信语音：微信、Viber。
- 零售行业：淘宝、Amazon。

目录

1. 计算在生活中的应用
2. 计算发展历程
 - 原始计算时代
 - 机械计算时代
 - 电子计算时代
3. 计算发展趋势

计算的起源

- 数学计算的发展始终伴随着人类的发展，原始社会时期的原始人类对于计算的诉求就是记录自己采摘果实的数据、狩猎动物的数量、氏族成员的数量，而这些对于计数的需求，促成了计算的诞生。
- 依据各种数学法则进行符号变换的过程就是计算，例如加减乘除就是最常见的数学运算法则，为了方便的进行运算，人类一直在寻找能够加快计算的方式方法，创造新的计算工具。



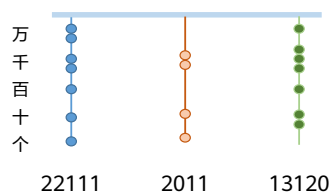
人类早期的计算工具

- 计算的演进史就是计算工具的演进史，而人们最早使用的计算工具就是自己的手指。在英语中“digit”一词同时具有数字，手指脚趾的意思。使用手指脚趾来计数显然是非常方便的，不过随着原始社会的发展，物质也变得逐渐丰富，这种简易的计算方式在应对超过10以上的计数时就力不从心了。在智人阶段，人们会使用石块来辅助计算，但这种方式依然算不上高效，仅仅只是稍微提高了些计算上限，一旦超过计数上限，也一样没有办法继续计算了。



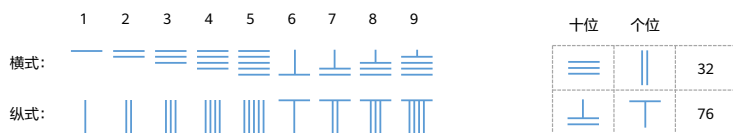
结绳计数

- 为了提高计数上限，古人们在绳子上面打结进行计数，这种方式可以较为长久的保存数量信息。在历史上有多个地区的不同种族的人类都曾使用过这一计数方式。例如在古代南美，印加人用各种动物的毛，例如羊驼或者骆马的毛来编织出颜色各异的绳子，名叫“基普”，这个名字由是古印加部落语言音译而来，含义为“打结的绳”。
- 基普绳分为主绳和副绳，副绳被垂直的系在较粗的主绳上，一般系有一百多根，有时甚至多达两千根。副绳上距离主绳不同的位置打着一个个结，印加人用结的位置距离主绳的远近来表示不同的数位，比如离主绳最远的结是个位，稍近一点的结是十位，最接近主绳的结代表万位，万也是印加人的最大计数单位。
- 印加人使用基普绳时，他们根据基普绳的颜色，结的数量，形状和位置，甚至是旋转方向来记录各种事件和自然现象，也被用在统计上，基普绳常被用于统计部落人口数目。



书契计数和算筹计数

- 书契计数出现在结绳计数方法之后，古人利用刻刀将数刻在兽骨、树木、龟甲、石壁等不容易被损坏的物体上，达到长久保存的目的。
- 公元前770年至公元前221年的春秋战国时期，人类的部落规模逐步扩大成为国家，刻字计数的方式也无法满足当前数据记录的需要了。中国人在此时使用算筹和算筹计数法来进行计算。算筹是多根具有相同长度与粗细的小细棍，可以使用多种材料进行制作，多用竹子制成。这些算筹可以被统一装在一个小袋子里随身携带，需要时就可以放在地上或者是桌上进行计算。
- 在算筹计数法中，用横纵两种排列方式来表示单位数目。表示多位数时，个位用纵式，十位用横式，百位再用纵式，千位再用横式，以此类推纵式横式在不同的数位上交替出现。



- 春秋战国时期的算筹和筹算制度已经比较成熟，这一点从《老子》一书中的：“善计者不用筹策”，就可以看出当时使用算筹进行计算已经是比较普遍了。
- 通过《汉书·律历志》中关于算筹形状大小的记载，我们可以得知西汉算筹是直径约为0.23厘米，长约14厘米的圆形竹棍。
- 通过《隋书·律历志》的记载，到了隋代，算筹已是三棱形与四棱形两种，以区别正数与负数。宽约为0.59厘米，长约8.85厘米。这表明从汉到隋，算筹从圆而方，由长变短，以便运用。

算盘

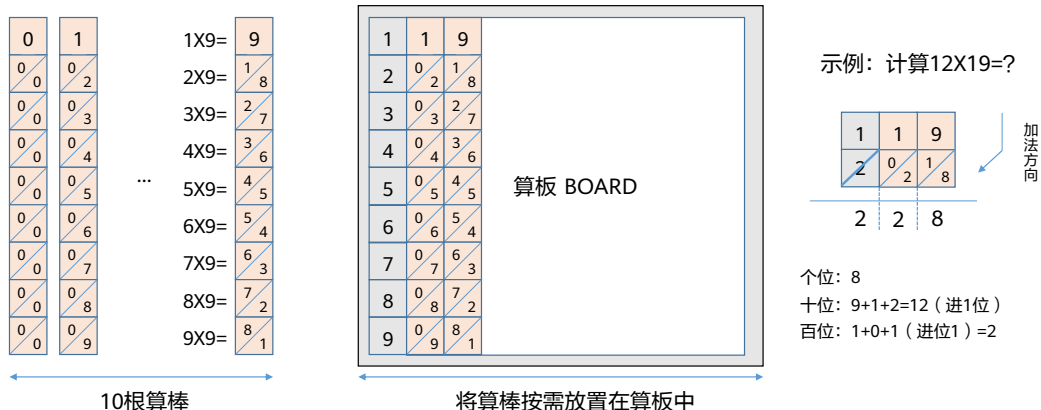
- 虽然算筹的出现是计算史上的里程碑事件，古人不仅可以利用算筹做简单的加减乘除的四则运算，还能够进行乘方和开方，但通过算筹进行计算也有其不足之处，由于需要通过摆放算筹进行计算，在计算量大时会耗费大量的时间进行摆放，非常麻烦。
- 为了解决这个问题，古人在算筹的基础之上发明了更好的工具——算盘。算盘是在阿拉伯数字出现之前曾被人们广为使用的一种计算工具。算盘制作容易，物美价廉，操作便捷，只要牢记珠算口诀后便可快速进行计算并得到结果。
- 中国算盘被传到世界各地，被不同国家的人们接纳和再创新，发展出了诸如俄罗斯算盘和日本算盘等在内的各式各样的形态的算盘。在电子计算时代的今天，算盘仍在一些特定的场合发挥着自己的作用。



- 算盘被很多学者认为是中华民族对世界的贡献之一，从清代开始就有许多学者对算盘是何时出现的这一问题展开了研究，然而由于缺少足够的史料，算盘究竟是何时出现的一直是众说纷纭。
- 随着人们进一步的研究，学者们认为算盘很有可能是起源于唐朝并在宋朝逐步发展成熟，这个观点的依据是在宋代的《清明上河图》中，一个算盘出现在了一家药铺的柜台上，这个算盘是与现代算盘的形状类似的串档算盘，可以推测在宋朝算盘的发展已经进入了成熟期。

纳皮尔筹

- 约翰·纳皮尔 (John Napier, 又译约翰·奈皮尔), 纳皮尔是苏格兰数学家, 他最伟大的贡献是发明了对数。1617年他出版了《筹学》, 并在此书中介绍了一种新形的手动计算工具, 也就是纳皮尔筹 (也被称作“纳皮尔骨”)。



- 纳皮尔筹的构造十分简单, 一般由一个算板和10根算棒组成。
- 在计算时算棒会被放置在算板中, 算板左侧边框标记着1-9, 与算棒上的方格一一对应。
- 每根算棒有10个方格, 顶部方格写有0-9的标志, 下方的9个方格都划有对角线, 分别填写着0-9与左边框对应数字的乘积, 并将十位数和个位数分开, 每根算棒都可根据计算需要来调整摆放的位置。
- 纳皮尔筹的意义是用加法和一位数乘法去替代多位数乘法, 从而简化运算难度, 同理它也可以用除法和减法代替多位数的除法。
- 这个工具其实和九九乘法表的作用很像。

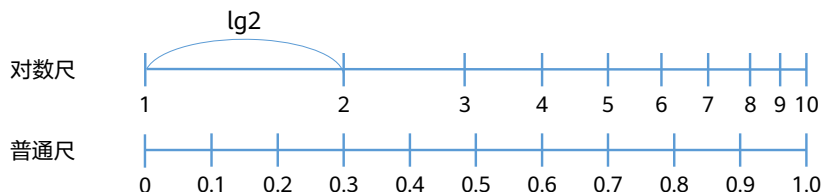
对数运算

- 16至17世纪，伴随着天文、工程、航海的迅猛发展，人们对于数字计算的需求变得越来越大，人们急需改进和优化数字计算方法进行应对。纳皮尔正是在研究天文学的过程中，为了寻求一种球面三角计算的简单方法时，运用他独特的数学方法和思路，构造了最初的对数。
- 对数原理在《奇妙的对数表的描述》这本书中被详细阐明，后人称为纳皮尔对数。
- 亨利·布里格斯是纳皮尔的朋友，他在研究了《奇妙的对数表的描述》一书后发现其中的对数使用起来很不方便，他建议纳皮尔将对数改良为以10为基底的对数表，由于人们常用的数系是10进制，因此会计算起来会更加方便。

$$\log_a X \longrightarrow \lg X$$

对数尺

- 根据计算对数的方法，科学家们发明了对数尺。对数尺可以将复杂的数值对应为刻度，这个刻度不同于不同尺子的“普通刻度”是等间距排列的，对数尺的使用“对数分度”的不等间距分度。



- 最初的对数尺使用起来还是比较复杂，受 $\log_a(MN) = \log_a(M) + \log_a(N)$ 和其他对数运算法则的启发，科学家们想到不如将两把对数尺组合起来形成一个新的计算工具，通过相对滑动来实现两把对数尺上读数的相加，根据这个思想，科学家们发明出了直线对数尺和圆形计算尺。

- 在之后的两百多年的时间里，计算尺也在不断发展，越来越多新的刻度被学者们引入到计算尺中，他们还为计算尺增加了可以移动的游标，让计算尺可以进行更为复杂的计算。
- 计算尺小巧方便便于携带，深受过去的工程师们喜爱，在二战期间常被用于计算导弹的飞行轨迹。

目录

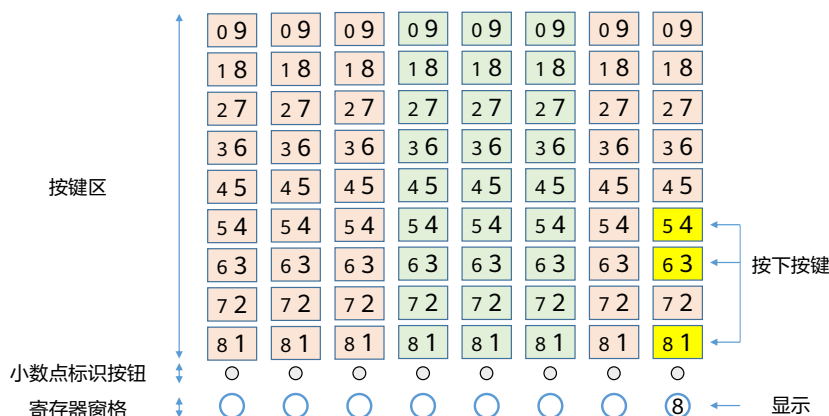
1. 计算在生活中的应用
2. 计算发展历程
 - 原始计算时代
 - 机械计算时代
 - 电子计算时代
3. 计算发展趋势

机械计算器——滚轮式加法器

- 滚轮式加法器，也称帕斯卡计算器，是由法国数学家布莱士·帕斯卡在他19岁时发明的，它通过转盘来输入数字进行加法运算，输出结果会显示在转盘上方的一个小窗口中。帕斯卡利用了一种巧妙的棘轮结构来解决计算器在把多个数字相加时会遇到的数字进位的问题，在计算过程中，内部的齿轮会进行连锁运作，例如当个位的齿轮正向旋转10位时，会使十位的齿轮对应增加一位。想要输入加数和被加数到计算器中时，转动计算器上的数字转盘到对应的刻度就行了，与拨电话号码盘非常相似。
- 早期的加法器还只能执行加法运算，无法直接在计算器上进行减法运算，随着技术的不断发展，后续滚轮式加法器可以实现通过负向的转动数字转盘来代表执行减法了。
- 滚轮式加法器的出现标志着人类的计算工具由手动向自动的演进，使用者无需理解加法器的运行原理，只需按照一定的规则操作加法器即可快速的获得运算结果。此外滚轮式加法器的机械结构及其设计思想也为后人带来了启发。

机械计算器——键驱机式计算器

- 1886年，美国工程师多尔·尤金·菲尔特发明了具有在计算史上具有重要意义的键驱式计算器，也称“菲尔特计算器”。在计算器上按下按键，就可以快速的得到结果。该计算器的造型与较早的打字机比较接近，在电子计算机发明前有着广泛的应用。



- 键驱机式计算器呈现方形盒状，下方有显示结果的寄存器窗格，盒子右侧有将窗格结果清空的手柄，每个按键上标有两个数，两个数之和为9。
- 按键上较大的数用于计算加法和乘法，每一列数与下方窗格代表的数位一一对应，每按一次键就会把该列按键的数值累加显示到窗格中，例如在同一列按下1、3、4，则下方显示为8。
- 按键上较小的数则用于减法和除法，例如计算500-427时，先按动500进入窗格中，再拨住百位窗格上方的按钮，同时依照按键上字号较小的那个数按动426（按动数需要比被减数小1），即可在窗格中得到计算结果73。这一过程实际上是利用补数，化减为加，为500加上了573，通过窗格上方的按钮阻止进位。除法则是减法的重复计算。

目录

1. 计算在生活中的应用
- 2. 计算发展历程**
 - 原始计算时代
 - 机械计算时代
 - 电子计算时代
3. 计算发展趋势

电子计算机发展简史



- 计算工具的演化经历了由简单到复杂、从低级到高级的不同阶段，例如从“结绳记事”中的绳结到算筹、算盘、计算尺、机械计算机等。它们在不同的历史时期发挥了各自的历史作用，同时也启发了现代电子计算机的研制思想。
- 1889年，美国科学家赫尔曼·何乐礼研制出以电力为基础的电动制表机，用以储存计算资料。
- 1930年，美国科学家范内瓦·布什造出世界上首台模拟电子计算机。
- 1946年，美国军方定制了世界上第一台电子计算机“电子数字积分计算机”。
- 1950年，第一台并行计算机问世，实现了计算机之父“冯·诺伊曼”的两个设想：采用二进制和存储程序。
- 1954年，IBM公司制造的第一台使用晶体管的计算机，增加了浮点运算，使计算能力有了很大提高。
- 1970年，IBM S/370，这是IBM的更新换代的重要产品，采用了大规模集成电路代替磁芯存储，小规模集成电路作为逻辑元件，并使用虚拟存储器技术，将硬件和软件分离开来，从而明确了软件的价值。
- 1975年，MITS制造世界上第一台微型计算机。
- 1977年，第一个带有彩色图形的个人计算机。
- 1998年，带有128K二级高速缓存的赛扬处理器成为广大装机者的最爱，同时64 M内存和15寸显示器开始成为标准配置。
- 1999年，部分品牌厂商开始将Pentium III CPU作为电脑的一个卖点，64 M内存和6.4 G硬盘开始成为电脑的标准配置。2001年至今，Pentium 4 CPU和Pentium 4赛扬CPU开始成为电脑的标准配置，内存由SDRAM实现了向DDR的过渡，同时17寸CRT显示器或者15寸液晶显示器开始成为用户的首选，硬盘逐渐向40 G以上的容量发展。

真空二极管

- 美国的知名科学家托马斯·爱迪生发明了碳丝电灯并投入使用后，一直在想办法对它进行优化，因为碳丝难以长时间承受电流产生的高温，不久后就会被蒸发，也就宣告着灯泡寿命的结束。在1883年他尝试将一根铜丝封入灯泡中，或许可以阻止灯丝被蒸发，而在反复试验之后，虽然这个方法没能解决灯丝被蒸发的问题，但是爱迪生观察到一个奇妙的现象，在碳丝温度不断升高后，与碳丝并不相接上的铜丝竟然有微弱的电流通过，爱迪生意识到这一现象可能会对后续人们研发电子设备有帮助，为此他先申请了专利，并将这个现象命名为“爱迪生效应”。
- 爱迪生效应的发现并未在当时对于爱迪生优化灯泡使用寿命的研究起到太大帮助，不过这个现象吸引了英国科学家约翰·安布罗斯·弗莱明的注意力，经过研究后，他认为在灯丝板极之间的空间是电的单行道。后来他在研究如何改进无线电报接收机的检波器时，尝试利用爱迪生效应进行检波。弗莱明把两个金属片封入真空玻璃管内，然后给阳极板加上高频交变电压后，出现了爱迪生效应，交流电在通过这个玻璃管时被转变成了直流电，弗莱明将这个装置称作真空二极管并在1904年为这个发明取得了专利。

真空三极管

- 真空二极管的出现是人类进入电子时代的里程碑事件，而真空三极管的出现使得电子管成为了被广泛应用于各类电子设备中的电子元器件。
- 1906年德·福雷斯特在复现弗莱明的实验时，尝试加入了第三根“Z”型，形状像是栅栏的电极，形成了电子管的第三个极，他发现只要把一个微小的变化电压加在第三极上，就可以在金属屏极上接受到一个电信号被放大的电流，这也意味着第三极对电流有控制作用，可以用于实现电信号的放大。德·福雷斯特把这个新的器件命名为真空三极管。由于真空三极管能够放大信号的特性，很快就被应用到无线通信领域中，在收音机和音响等设备上都得到了广泛的应用。
- 真空三极管不仅可以实现电流的放大，还可以实现开关的作用，例如使得电流在阴极到屏极之间流通，或者是在阴极和屏极之间没有电流，状态的变化可以通过调整第三极上的电压来实现。真空三极管充当开关器件的性能要远强于继电器设备，使得大幅提升计算器计算速度成为了可能。
- 真空三极管为电子计算机的诞生创造了先决条件，后续人们还进一步研制出了四极和五极的真空管，人们把他们都统称为电子管。在第一台电子计算机ENIAC里就包含了17000多根电子管。

ENIAC

- ENIAC (Electronic Numerical Integrator And Computer, 电子数字积分计算机), 它是第二台电子计算机, 也是第一台通用计算机。研制一台计算性能强大的电子计算机的想法产生于第二次世界大战期间, 飞机和大炮在当时战场上的地位是举足轻重的, 美国军方设立了弹道研究室以研制和开发新型武器装备, 当时研究员们每天都被要求计算几百条弹道来为军方提供射表, 但巨大的计算压力使之成为了一个不可能完成的任务。在和时间赛跑的战争年代, 显然人们需要研发出一台高效的计算机来摆脱这个困境, 于是在当时成立了一个以任职于宾夕法尼亚大学莫尔电机工程学院的莫希利为首的研究小组, 在冯·诺依曼等人的参与下, 在1946年的美国宾夕法尼亚大学, ENIAC被成功研发。
- ENIAC能够进行多达每秒5000次左右的加法运算, 每秒400次左右的乘法运算, 这样的计算能力已经比当时已有的传统计算机快了近千倍, 也可以通过再编程来解决一些更为复杂的函数计算。这极大的提高了弹道计算效率, 填上了算力极度不足的缺口, 开启了电子计算机的新时代。

- 阿塔纳索夫-贝瑞计算机, 也称为ABC计算机, 是世界上第一台电子计算机, 由美国科学家阿塔纳索夫在1937年设计。但该电脑的功能十分有限, 也无法编程, 它的设计的目的仅为求解线性方程组。
- ENIAC能够再编程, 满足解决各种各样计算问题的需求。这台计算机是个庞然大物, 它长30米多, 宽6米, 高约2米多, 重30多吨。在它的内部含有17000多根电子管, 7000多根水晶二极管, 以及其他数量繁多的电阻、电容、继电器等等电子元器件。

晶体管与第二代电子计算机

- 在第二次世界大战结束后，虽然真空管已经被广泛的应用到电子学研究中，但是它体积较大，可靠性低，很容易由于过热而产生故障的种种缺点却一直困扰着科学家们，为了维修或更换损坏的真空管，耗费的时间和人力成本是巨大的。贝尔实验室的威廉·肖克利希望使用半导体来替代真空管以解决这些问题。
- 于是在1947年，肖克利与同在贝尔实验室的约翰·巴丁、沃尔特·布拉顿三人成功的制造出第一个锗晶体管。晶体管的出现让计算机的计算能力再次飞跃，它不仅可以实现电子管的功能，还具有体积小，重量轻，发热低的特点，1954年贝尔实验室研制出了第一台使用晶体管线路的计算机TRADIC，它装有约800只晶体管，功耗也只有100瓦，而在4年之后IBM公司制成了全部使用晶体管的RCA501型计算机，得益于第二代计算机的晶体管逻辑元件和新型存储器，计算速度从几千次每秒，一下提升到几十万次每秒，存储容量也提升至原来的百倍。

- 在肖克利研究晶体管的初期，他尝试使用硅材料去制作电子管，但受制于当时硅的提纯工艺的影响，无法获得具有良好导电效果的硅，因此巴丁和布拉顿两人在发现这个问题之后转而寻求使用锗元素去制作电子管，在他们初期的工作成果被肖克利进一步改进之后，终于制作出了具有实用性的晶体管。
- 不过锗材料的自然界储量远不如硅那么丰富，价格也比硅昂贵。由于后续材料提纯工艺的改进，在导电性良好的硅可以容易被人们获取后，硅晶体管逐步取代了锗晶体管的地位。

集成电路

- 1958年，当时还是一名德州仪器公司的新员工杰克·基尔比正待在冷清的公司，8月的德州十分炎热，此时他的同事们都去享受公司提供的高温假期去了，周围的环境噪声少了许多，让基尔比可以专心的工作和思考。当前公司有非常多的技术工人每天需要穿上不便活动的防护服再盯着显微镜去完成微型晶体管的焊接工作，在耗费大量时间的同时还很容易出错，导致计算机出现故障。基尔比正在研究如何微小化各类电子元器件，在安静的思考环境中，一个灵感涌现出来，既然电阻器和电容器的制造材料与晶体管相同，与其将它们分别制造出来在一个个相互焊接在一起，何不把这些元件都先在同一块半导体材料上制造出来，然后就地直接在同样一块半导体上制作所需的电路呢？
- 有了这个绝妙的想法后，他将思路整合后记录下来，写了一些材料和主管和同事们讨论，大家很快都理解了这个设计的重要意义并开始了测试研究，没过太久，线路测试成功，德州仪器很快便宣布他们发明了集成电路，基尔比也为此申请了专利，不过此时的集成电路还并不完善，还无法被大规模使用。1959年仙童半导体公司的罗伯特·诺伊斯利用新发明的平面处理技术，使用了和德州仪器不一样的另一种方法制作出了以硅作为主要材料的集成电路，并投入了商业领域，集成电路才开始渐渐进入到人们的视野中，这也使得他与基尔比共享了集成电路之父的称号。而到现在，人们最为熟知的集成电路的典型代表，就是芯片了。

微处理器与Altair 8800

- 随着集成电路的不断发展，工程师们追求制造出更加高度集成的，体积更小的处理器芯片的脚步从未停歇。英特尔工程师特德·霍夫在1971年与他的同事们一起研发出了世界上首枚微处理器intel 4004，它集成了2000多个晶体管，能够处理4比特的数据，每秒能进行超过5万次的运算，但尺寸却只有12平方厘米，售价也并不昂贵，收到了市场的热烈追捧。
- Intel 4004珠玉在前，英特尔便开始持续改进微处理器并推出了包含Intel 8080在内的多款微处理器，这些处理器的诞生进一步帮助英特尔巩固了它的市场地位。
- 随后在1974年，MITS公司基于Intel 8080设计了世界上第一款微型电子计算机，名为Altair 8800，是世界上第一种面向个人使用并出售的电脑。它的出现立刻引起了电子设备爱好者的关注并很快取得了商业成功。

- Altair 8800的外形是一个长方形的金属盒子，没有显示器，键盘和鼠标等现在看来十分常见的外设，盒子的正面有几排小的指示灯来表示当前电脑工作状态，此时它还仅能吸引那些关注电子设备领域发展的专业人士或者是爱好者，普通的人们在没有技术基础的前提下是不容易上手使用的。

图形化计算机

- 1977年，Apple II上市，是全球首台8bit彩色图形计算机，搭载了MOS 6502微处理器，有4KB内存，支持BASIC编程语言。得益于简单易上手的图形化界面，作为非电子设备爱好者的普通人也能够以较低的学习成本使用计算机了，这很快使得Apple II成为了世界上第一台被成功批量生产并销售的家用计算机。
- 到了1983年，Apple Lisa上市，成为了全球首台同时使用图形化界面和鼠标的个人电脑，但这台家用电脑在当时由于其昂贵的售价并没有取得太大的成功，好在Apple Lisa上很多优秀的软件设计被继续沿用至后来的Macintosh电脑上，而Macintosh凭借着其简洁美观的外形设计，对用户使用友好的漂亮的图形界面，再一次加速了个人家用电脑的普及。

超级计算机

- 时间回到21世纪，此时个人电脑已经被广泛普及，成为了每个人工作生活中不可或缺的一部分，小巧轻便的笔记本电脑里就可以装下性能强劲的CPU和GPU，拥有超快读写速度的大容量内存和固态硬盘，方便与用户交互的触控板、鼠标、屏幕。但人类对于强大计算能力的渴望不会止步于此，于是超级计算机应运而生。
- 超级计算机，能够以超强的运算速度，超大的存储空间，更丰富的软硬件功能来处理个人计算机无法复杂计算数据。以中国“神威·太湖之光”超级计算机为例，它的持续计算性能能达到9亿亿次每秒，峰值性能更是能超过12亿亿次每秒。
- 超级计算机根据其使用场景可以大致分为两类，一类能快速的处理专一领域的计算工作，例如处理天体物理学的计算和密码学计算，另一类则使用非专用系统，服务于很多通用场景，例如气象，金融，医疗等领域。

- 超级计算机的硬件组成种类和一般的家用计算机大同小异，但是各个组件的规格却有着天壤之别。

目录

1. 计算在生活中的应用
2. 计算发展历程
- 3. 计算发展趋势**
 - 云计算简介
 - 大数据简介
 - 人工智能简介
 - 发展挑战与前景

计算产业概述

- 计算产业涵盖了与计算技术相关的广泛领域，例如个人电脑、网络服务器、软件中间件、数据存储解决方案等等信息技术，它是与信息技术相关的各种产业的统称。
- 计算产业是IT技术的基础，是每一次产业变革的驱动力，从云计算、大数据、人工智能再到区块链、边缘计算、物联网等等高新技术都离不开强大的计算能力的支持。本章将以云计算、大数据和人工智能作为计算产业的代表，通过对它们的介绍来了解计算产业的发展和现状。



云计算



大数据



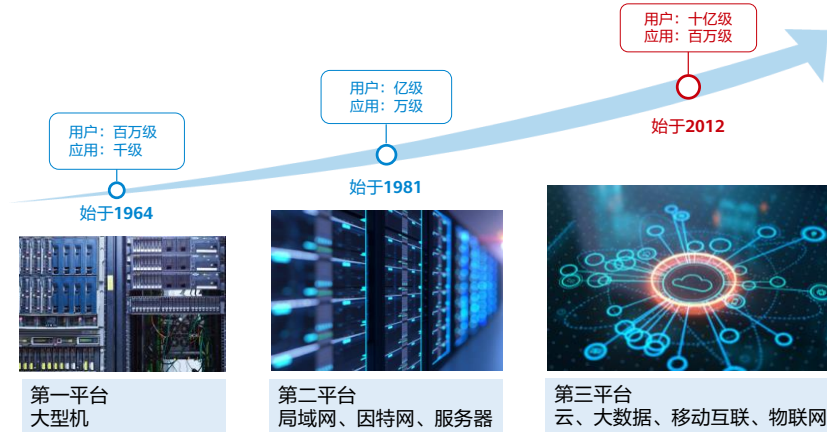
人工智能

什么是云计算

- 美国国家标准与技术研究院（NIST）定义
 - 云计算是一种模型，它可以实现随时随地，便捷地，按需应变地从可配置计算资源共享池中获取所需的资源（例如，网络、服务器、存储、应用及服务），资源能够快速供应并释放，使管理资源的工作量和与服务提供商的交互减小到最低限度。
- 维基百科定义
 - 云计算是一种通过Internet，以服务的方式提供动态可伸缩的虚拟化的资源的计算模式。



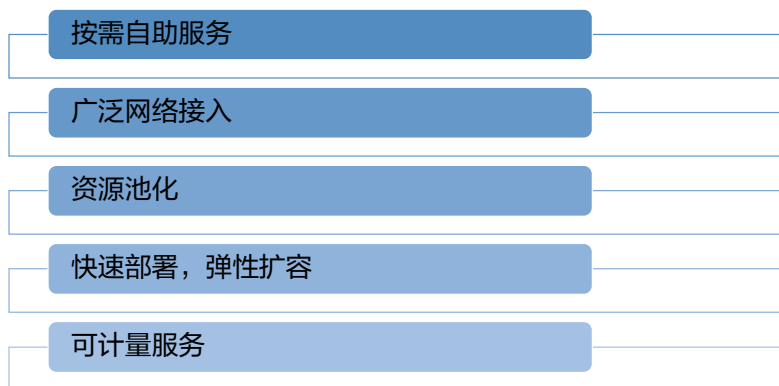
以云计算为基础的“第三平台”已经成为IT产业主流



- IDC的数据显示，2022年中国在ICT领域的投资超过了5300亿美元。IDC进一步预测到2027年，由于中国数字化转型的不断发展和创新，中国在ICT领域的总投资将达到7200亿美元以上，届时将占全球市场的11.7%，并且预期在这五年间，年均复合增长率将保持在6.2%左右。

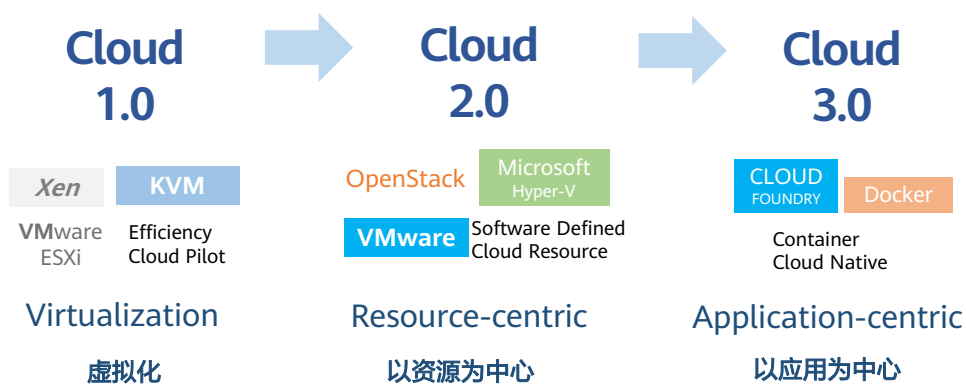
- 2015年，第二平台陷入衰退，第三平台不再是概念。
- 全球IT Spending的1/3来自第三平台，IT Spending增长100%由第三平台。
- 云计算改变了IT产业的商业模式和建设模式，大数据帮助企业挖掘商业价值和促进企业进行第二数据平面的建设。

云计算的价值



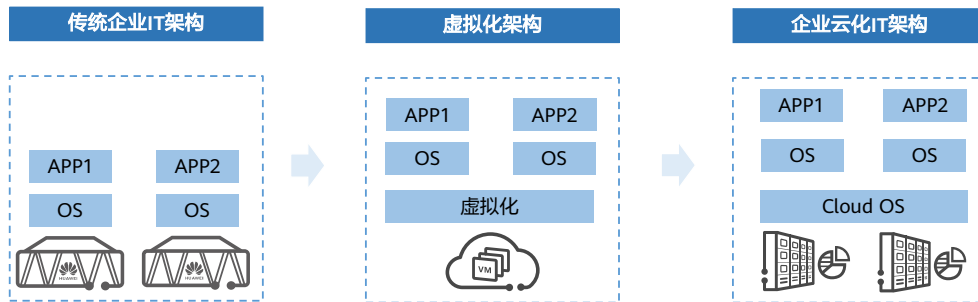
- 云计算通过将硬件资源以软件的方式整合为一个整体, 然后再以软件的方式动态分配给应用, 大大地提高了资源的使用率, 并且还能够弹性扩容, 极大地优化了工作效率。通过建设高规格的云数据中心, 引入自动化调度技术, 让数据存储更加集中, 数据资产也就能够更加有效利用, 也更加节能减排和易于维护。从各个维度都起到了降本增效的作用。
- 五个主要价值点:
 - 按需自助服务: 消费者可以按需部署处理能力, 如服务器时间和网络存储, 而不需要与每个服务供应商进行人工交互。
 - 广泛网络接入: 可以通过互联网获取各种能力, 并可以通过标准方式访问, 通过各种客户端接入使用。例如移动电话, 笔记本电脑, PAD等。
 - 资源池化: 供应商的计算资源被集中起来, 以便以多用户租用模式服务所有客户, 同时不同的物理和虚拟资源可根据客户需求动态分配和重新分配。客户一般无法控制或知道资源的确切位置。这些资源包括存储、处理器、内存、网络带宽和虚拟机等。
 - 快速部署, 弹性扩容: 云计算可以迅速、弹性地提供能力, 能快速扩展, 也可以快速释放实现快速缩小。对客户来说, 可以租用的资源看起来似乎是无限的, 并且可在任何时间购买任何数量的资源。
 - 可计量服务: 云服务的收费是基于用户实际使用的资源进行计量, 比如云主机的CPU、内存、存储容量, 网络带宽的消耗, 按小时计费或者包年包月。

云计算的发展



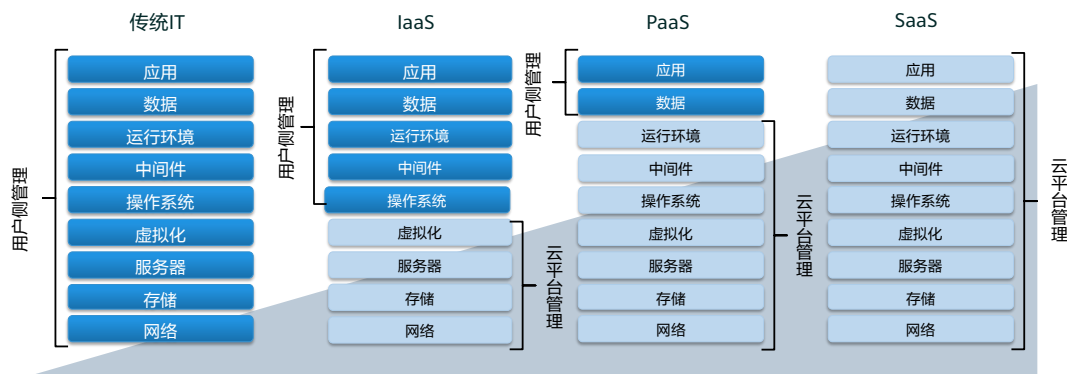
- 针对上述三大云计算发展演进里程碑阶段而言，云计算1.0普遍已经是过去式。且一部分行业、企业客户已完成初步规模的云计算2.0建设商用，正在考虑该阶段的进一步扩容，以及面向云计算3.0的演进；而另一部分客户则正在从云计算1.0走向云计算2.0，甚至同步展开云计算2.0和3.0的演进评估与实施。
- Cloud 1.0：强调虚拟化，追求更高的资源利用率。
- Cloud 2.0：以资源为中心，实现基础设施云化、资源服务标准化、自动化。
- Cloud 3.0：以应用为中心，实现应用云化，敏捷应用开发与生命周期管理。

企业IT基础设施架构开始走向云化



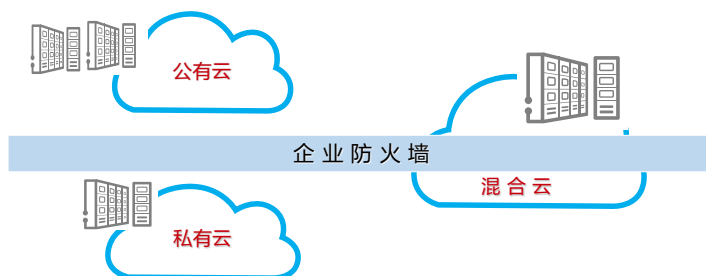
- 传统IT基础架构由通常的硬件和软件组件组成：设施、数据中心、服务器、网络硬件、台式计算机和企业应用软件解决方案。与其他基础架构类型相比，这种基础架构设置通常需要更多的电力、物理空间和资金。传统基础架构往往安装在本地，仅供企业或专有使用。
- 虚拟化是指计算机元件在虚拟的基础上而不是真实的基础上运行。虚拟化技术可以扩大硬件的容量，简化软件的重新配置过程。
- 企业数据中心“云化”转型的要点：
 - 从资源孤岛到真正资源池化。
 - 从集中式向分布式架构转型。
 - 从专用硬件向开放的软件定义模式转型。
 - 从人工处理向自助、自动服务转型。
 - 从分散统计到统一计量转型。

云计算的服务模式



- IaaS: Infrastructure as a Service, 基础设施即服务, 就是由云平台提供基础设施 (如: 服务器、存储、网络、虚拟化资源) 以及负责相关资源的维护, 用户只需要关注系统和应用层面的部分即可。
- PaaS: Platform as a Service, 平台即服务, 就是由云平台提供基础设施 (如: 服务器、存储、网络、虚拟化资源) + 应用部署环境 (如: 操作系统、中间件、软件运行环境) 以及负责相关资源的维护, 用户只需要关注应用和数据本身即可。
- SaaS: Software as a Service, 软件即服务, 就是由云平台提供全部资源服务以及维护, 用户只管使用应用即可。
- 相比于传统IT全流程全设备采购的方式, 云服务模式将IT设备服务化地销售, 让客户按需选择, 在使用灵活性和成本上比传统IT更有优势。

云计算的部署模式



私有云 (Private cloud) : 云计算的基础设施由单一的组织拥有, 并且仅仅为该组织运营。

公有云 (Public cloud) : 云服务运营商拥有云基础设施, 并且为公众或者企业用户提供云服务。云计算基础设施由一个组织拥有并且向公众或者大型的工业团体销售云计算服务。

混合云 (Hybrid cloud) : 云计算基础设施由两种或者多种云组成, 对外仍然表现为一个整体。

- **私有云:** 私有云通常部署在企业或单位内部, 运行在私有云上的数据全部保存在企业自有的数据中心内, 如果需要访问这些数据, 就需要经过部署在数据中心入口的防火墙, 这样可以在最大程度上保护数据。
- **公有云:** 云服务运营商拥有云基础设施, 并且为公众或者企业用户提供云服务。云计算基础设施由一个组织拥有并且向公众或者大型的工业团体销售云计算服务, 用户可以通过互联网像使用水电一样使用IT服务。
- **混合云:** 混合云是一种比较灵活的云计算模式, 它可能包含了公有云、私有云或者后面要讲的行业云中的两种或两种以上的云, 用户的业务可以根据需求在这几种云上切换。

目录

1. 计算在生活中的应用
2. 计算发展历程
- 3. 计算发展趋势**
 - 云计算简介
 - 大数据简介
 - 人工智能简介
 - 发展挑战与前景

什么是大数据

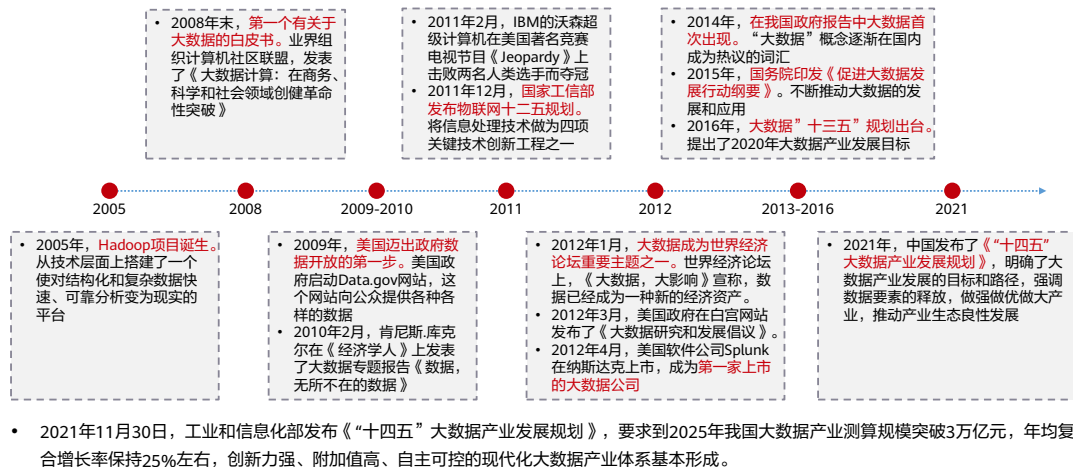
- 维基百科：由巨型数据集组成，这些数据集大小常超出人类在可接受时间下的收集、使用、管理和处理能力。大数据也可以定义为来自各种来源的大量非结构化和结构化数据。
- 结构化数据：可以使用关系型数据库表示和存储，表现为二维形式的数据。一般特点是，数据以行为单位，一行数据表示一个实体的信息，每一行数据的属性是相同的。
- 半结构化数据：半结构化数据是结构化数据的一种形式，它并不符合关系型数据库或其他数据表的形式关联起来的数据模型结构，但包含相关标记，用来分隔语义元素以及对记录和字段进行分层。常见的有XML和JSON。
- 非结构化数据：没有固定结构的数据。各种文档、图片、视频音频等都属于非结构化数据。对于这类数据，我们一般直接整体进行存储，而且一般存储为二进制的格式。

id	name	age	gender
1	Tim	12	male
2	Harry	14	male
3	Ann	18	female

```
<person>
  <name>A</name>
  <age>13</age>
  <gender>female</gender>
</person>

<person>
  <name>B</name>
  <gender>male</gender>
</person>
```

大数据发展重要历程



- 2014年3月，大数据首次写入中国中央政府工作报告。
- 2015年10月，党的十八届五中全会正式提出“实施国家大数据战略，推进数据资源开放共享”。这表明中国已将大数据视作战略资源并上升为国家战略，期望运用大数据推动经济发展、完善社会治理、提升政府服务和监管能力。
- 2017年12月8日，习近平主席在中共中央政治局第二次集体学习时强调，审时度势精心谋划超前布局力争主动，实施国家大数据战略加快建设数字中国。
- 2018年5月，习近平在向中国国际大数据产业博览会的致辞中指出，我们秉持创新、协调、绿色、开放、共享的发展理念，围绕建设网络强国、数字中国、智慧社会，全面实施国家大数据战略，助力中国经济从高速增长转向高质量发展。
- 2020年5月工信部颁发《关于工业大数据发展的指导意见》，推动工业数据全面采集，加快工业设备互联互通，推动工业数据高质量汇聚，统筹建设国家工业大数据平台，推动工业数据开放共享，激发工业数据市场。

大数据实现数据增值



一个大规模生产、分享和应用数据的时代正在开启.....

——肯尼思·库克耶《大数据时代》作者

- 大数据技术的战略意义不在于掌握庞大的数据信息，而在于对这些含有意义的数据进行专业化处理。换言之，如果把大数据比作一种产业，那么这种产业实现盈利的关键，在于提高对数据的“加工能力”，通过“加工”实现数据的“增值”。
- 从技术上看，大数据与云计算的关系就像一枚硬币的正反面一样密不可分。大数据必然无法用单台的计算机进行处理，必须采用分布式架构。它的特色在于对海量数据进行分布式数据挖掘。但它必须依托云计算的分布式处理、分布式数据库和云存储、虚拟化技术。
- 随着云时代的来临，大数据（Big Data）也吸引了越来越多的关注。大数据（Big Data）通常用来形容一个公司创造的大量非结构化数据和半结构化数据，这些数据在下载至关系型数据库用于分析时会花费过多时间和金钱。大数据分析常和云计算联系在一起，因为实时的大型数据集分析需要像MapReduce一样的框架来向数十、数百或甚至数千的电脑分配工作。

大数据处理与传统数据处理的差异

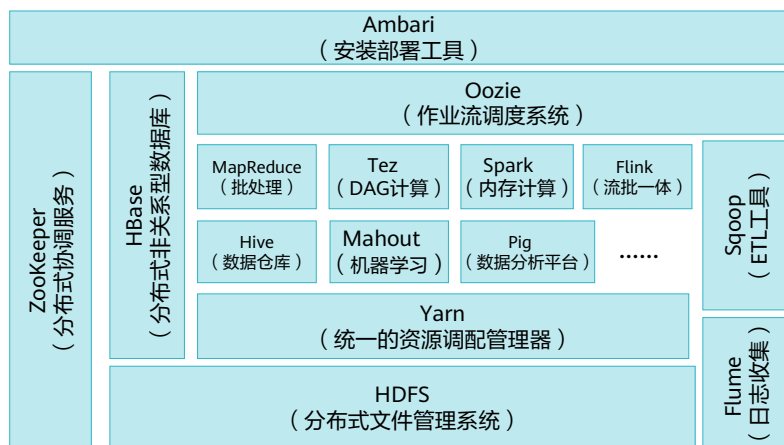
- 从数据库（Database，DB）到大数据（Big Data，BD）。
 - “池塘捕鱼” VS “大海捕鱼”，“鱼”是待处理的数据。

	传统数据处理	大数据处理
数据规模	小（以MB为处理单位）	大（以GB、TB、PB为处理单位）
数据类型	单一（结构化为主）	繁多（结构化、半结构化、非结构化）
模式和数据的关系	先有模式后有数据（先有池塘后有鱼）	先有数据后有模式，模式随数据增多不断演变
处理对象	“池塘中的鱼”	“大海中的鱼”，通过某些“鱼”判断其他种类的“鱼”是否存在
处理工具	One size fits all	No size fits all

大数据技术发展



Hadoop大数据生态圈



大数据应用的主要计算模式

- 批处理计算
 - 针对大规模数据的批量处理。主要技术有MapReduce、Spark等。
- 流计算
 - 针对流数据的实时计算处理。主要技术有Spark、Storm、Flink、Flume、Dstream等。
- 图计算
 - 针对大规模图结构数据的处理。主要技术有GraphX、Gelly、Giraph、PowerGraph等。
- 查询分析计算
 - 大规模数据的存储管理和查询分析。主要技术有Hive、Impala、Dremel、Cassandra等。

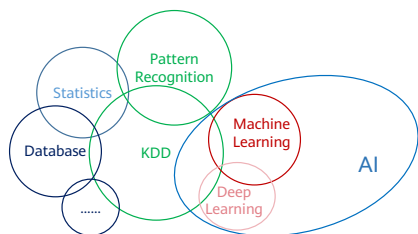
- GraphX: Spark内置的图计算算法库。
- Gelly: Flink内置图计算算法库。
- Giraph: Google 于 2010 年发布的论文Pregel: a system for large-scale graph processing的开源实现。Giraph 是以 Hadoop 为基础开发的上层应用，其系统架构和计算模型与 Pregel 保持了一致。
- PowerGraph: 分布式图并行计算框架，为了解决Natural Graph中的“长尾现象”和“幂律分布”问题而提出了节点切分策略。
- Impala是Cloudera公司主导开发的新型查询系统，它提供SQL语义，能查询存储在Hadoop的HDFS和HBase中的PB级大数据。相比于Hive，速度更快。
- Dremel: 谷歌的交互式数据分析系统，同样能提供比Hive更快的速度，可作为Hive和MapReduce的补充。被Impala借鉴。
- Cassandra是一套开源分布式NoSQL数据库系统。它最初由Facebook开发，用于储存收件箱等简单格式数据，集GoogleBigTable的数据模型与Amazon Dynamo的完全分布式的架构于一身。

目录

1. 计算在生活中的应用
2. 计算发展历程
- 3. 计算发展趋势**
 - 云计算简介
 - 大数据简介
 - 人工智能简介
 - 发展挑战与前景

什么是人工智能

- 人工智能（Artificial Intelligence）的定义可以分为两部分，即“人工”和“智能”。“人工”即由人设计，为人创造、制造。
- 人工智能它是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学。1956年由约翰·麦卡锡首次提出，当时的定义为“制造智能机器的科学与工程”。人工智能的目的就是让机器能够像人一样思考，让机器拥有智能。
- 时至今日，人工智能的内涵已经大大扩展，是一门交叉学科。



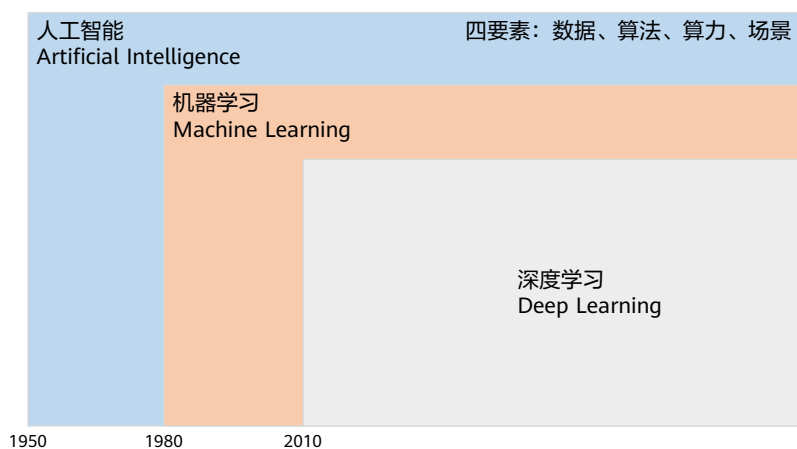
47 Huawei Confidential



HUAWEI

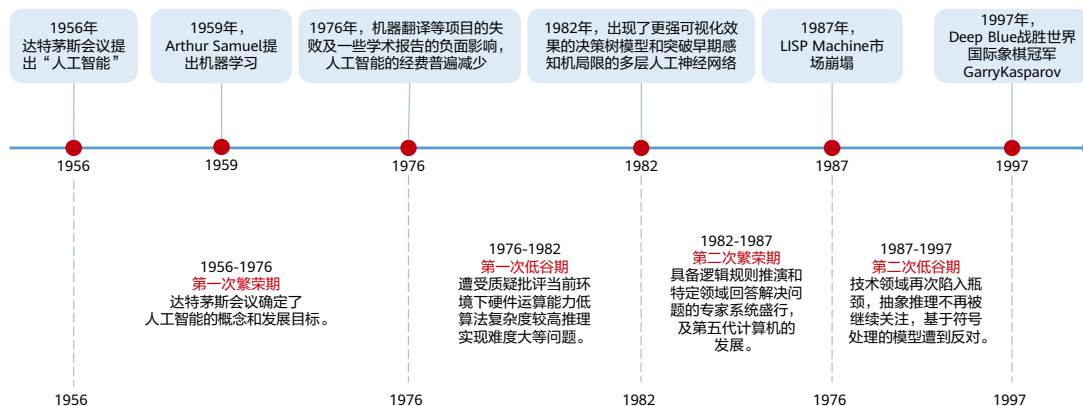
- 对机器学习的认识可以从多个方面进行，有着“全球机器学习教父”之称的 Tom Mitchell则将机器学习定义为：对于某类任务T和性能度量P,如果计算机 程序在T上以P衡量的性能随着经验E而自我完善，就称这个计算机程序从经验E学习。这些定义都比较简单抽象，但是随着对机器学习了解的深入，我们会发现随着时间的变迁，机器学习的内涵和外延在不断的变化。因为涉及到的领域和应用很广，发展和变化也相当迅速，简单明了地给出“机器学习”这一概念的定义并不是那么容易。
- 普遍认为，机器学习（Machine Learning，常简称为ML）的处理系统和算法是通过找出数据里隐藏的模式进而做出预测的识别模式，它是人工智能（Artificial Intelligence，常简称为AI）的一个重要子领域，而人工智能又与更广泛的数据挖掘（Data Mining，常简称为DM）和知识发现（Knowledge Discovery in Database，常简称为KDD）领域相交叉。上图可以更好的帮助我们理解人工智能（Artificial Intelligence）、机器学习（Machine Learning）、数据挖掘（Data Mining）、模式识别（Pattern Recognition）、统计（Statistics）、神经计算（Neuro Computing）等概念之间的关系。

AI、机器学习、深度学习的关系

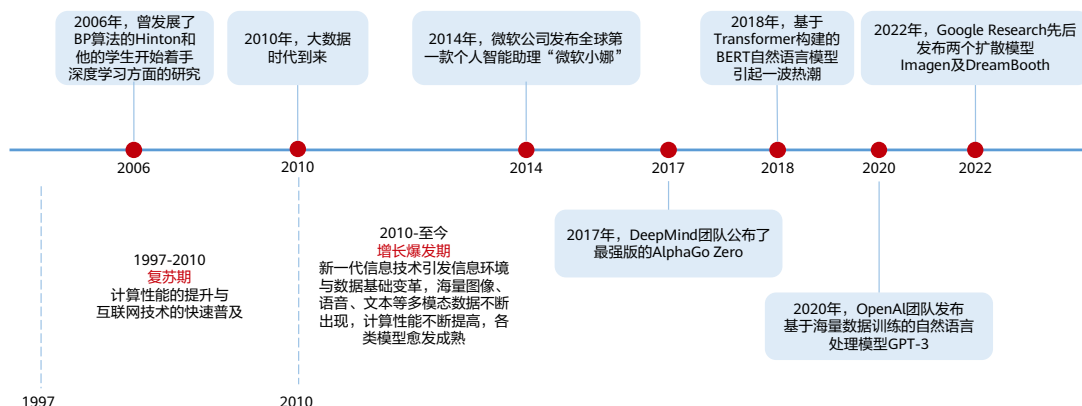


- 人工智能：是研究、开发用于模拟、延伸和扩展人的智能的理论、方法及应用系统的一门新的技术科学。“人工智能”一词以前曾被用来描述模仿和展示与人类思维相关的“人类”认知技能的机器。
- 机器学习：专门研究计算机怎样模拟或实现人类的学习行为，以获取新的知识或技能，重新组织已有的知识结构使之不断改善自身的性能。机器学习算法主要是基于样本数据（称为训练数据）构建模型，以便能在没有明确编程的情况下做出预测或决策。
- 深度学习：深度学习的概念源于人工神经网络的研究。多层感知机（MLP，Multilayer Perceptron）就是一种深度学习结构。深度学习通过组合低层特征形成更加抽象的高层表示属性类别或特征，以发现数据的分布式特征表示。深度学习是机器学习研究中的一个领域，它模仿人脑的机制来解释数据，例如图像，声音和文本。

人工智能发展简史（1）

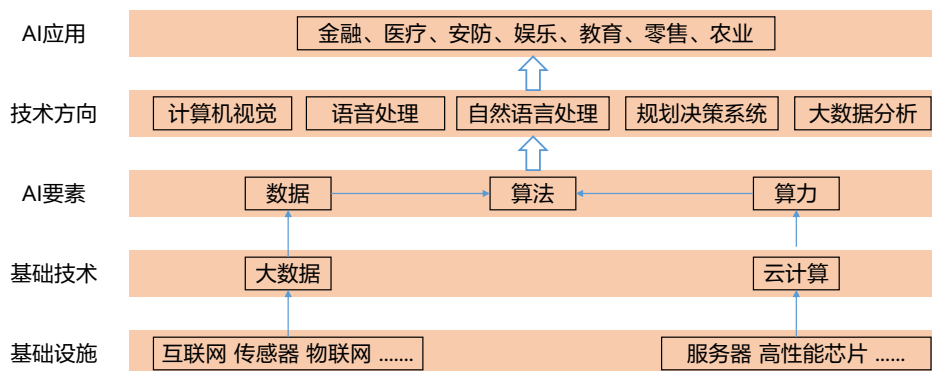


人工智能发展简史（2）

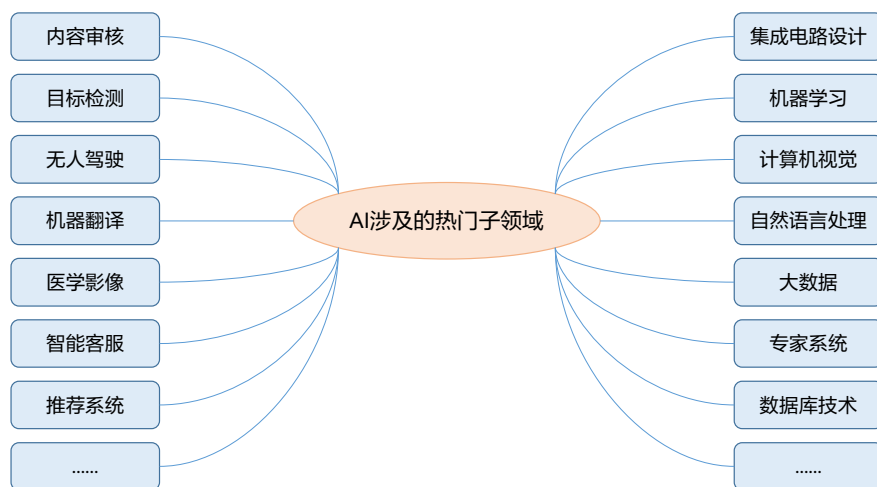


AI产业生态

- 人工智能应用落地的四要素是数据、算法、算力、场景。为实现智能社会，我们需要将AI与云计算、大数据和物联网等结合。

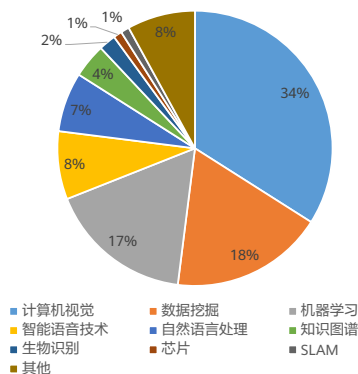


AI热门子领域



国内人工智能企业应用技术分布

- 中国人工智能企业的十大应用技术领域中，计算机视觉、机器人、自然语言处理、机器学习、生物识别占比居前五。
- 目前AI的通用技术方向主要为：
 - 计算机视觉：
 - 是研究如何让计算机“看”并“看”地更快更精确。
 - 自然语言处理：
 - 是研究让计算机来理解并运用自然语言。



REF: 亿欧智库《2022中国人工智能芯片行业研究报告》

- 技术层：通用技术，算法，技术框架。
- 通用技术：自然语言处理，计算机视觉，语音识别，机器学习等。
- 算法：机器学习，强化学习，深度学习等。
- 技术框架：分布式存储，分布式计算等。

目录

1. 计算在生活中的应用
2. 计算发展历程
- 3. 计算发展趋势**
 - 云计算简介
 - 大数据简介
 - 人工智能简介
 - 发展挑战与前景

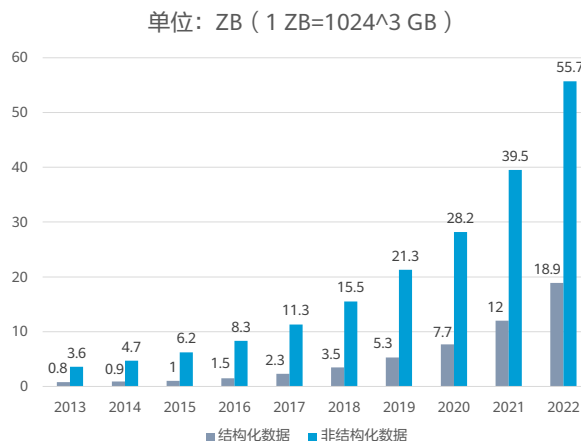
大数据技术对算力带来挑战



- 大数据1.0面向程序员， 2.0面向分析师， 3.0 面向普通工程师。
- 在互联网时代，相比传统的企业电子办公，出现了大量的非结构化数据，对数据的分布式存储和并行技术提出了很高的要求。Google就是传统互联网的典型代表，发表了GFS、MR、BigTable等重要的大数据论文，从而引导了大数据的元年。
- 在移动互联网时代，在互联网上的应用远远比传统互联网（PC访问）丰富，数据的结构也发生了很大的变化，对实时处理的能力和大量并发的要求比以前会有巨大的提高。手机银行就是一个典型的代表，现在的手机银行APP能够提供各种丰富的内容，除了传统的网上交易，还包括资讯，娱乐，购物等。所以对数据处理的能力通常就要求多种技术配合来完成，有批处理，内存计算，交互式查询，流式实时处理等，这样就推动了MR, SPARK, Solr, Storm和YARN的发展。
- 到了即将爆发的物联网时代，将是一个完全的数字化的世界，任何东西都将连接到网络上，而最关键的是要将这些东西关联起来，发现其背后存在的规律，为人类的生活、生产创造更大的价值，这就是人工智能时代的大数据处理，这是大数据发展的第3个阶段。

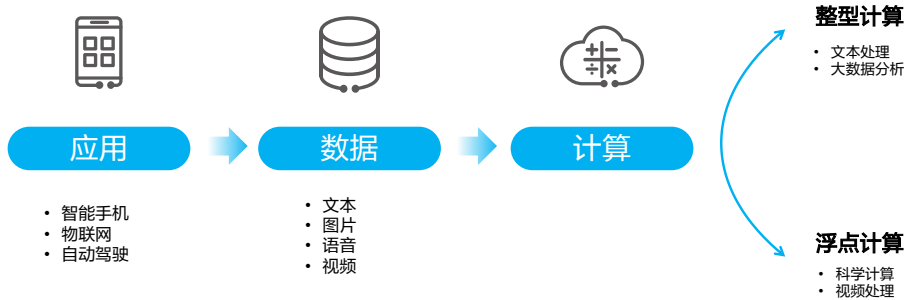
人工智能对算力带来挑战

- 数据爆发式增长带来算力需求提升。
- 企业资源投入以及算力成本不断增加。
- 摩尔定律失效，CPU性能提升遭遇瓶颈，传统服务器无法满足算力需求。
- AI计算系统性能与效率亟待提升。



- 目前的人工智能，更多的是代表智能的个体，能够通过自身的持续学习能力，智能的完成单点决策。
- AI产业爆发三大条件：算法、算力、数据。
- 机器经验需要由大量的历史数据得出，所以数据收集无处不在，数据的增长会是几何级数的。使用当前集中式的存储和集中式的通信模式，在未来是无论如何都无法通过一个巨型单点支撑如此大的体量，存储和通信能力都是瓶颈，而且效率会非常低下。
- 算力成本也是人工智能行业的一大痛点。现在的人工智能企业的硬件投入巨大。人工智能对计算的需求非常大，因此对高性能计算定制深度学习芯片要求很高，意味着很多企业要花很多钱买算力、建很多计算中心，造成了很大的资源浪费。
- AI计算发展趋势演变过程中面临着巨大的挑战：随着模型所需的精度越高，所需的计算量也会呈现增长趋势。对于未来算法的发展对整个计算需求所造成的挑战会变得更加，提高整个AI计算系统的性能与效率显得尤为重要。
- 摩尔定律失效，CPU性能提升遭遇瓶颈。Intel宣布正式停用“Tick-Tock”处理器研发模式，未来研发周期将从两年周期向三年期转变。单颗CPU性能的提升在放缓，传统服务器难以满足并行算力需求，服务器CPU出货量增长停滞。
- 如今，AI面临着巨大的计算挑战，提高AI计算系统性能与效率变得尤为重要，需要从系统的角度进行综合考虑。

应用和数据的多样性需要新的计算架构



- 未来信息量巨大，计算无处不在，计算应用的场景多种多样，从扫地机器人，到智能手机、智慧家庭、IoT物联网、智能驾驶等。场景的多样性，带来数据的多样性，数字、文本、图片、视频、图像以及结构化数据、非结构化数据等。
- 整型计算胜在文本处理、存储、大数据等；浮点计算胜在科学计算等。

新兴计算产业呼唤新的算力

移动智能终端逐渐取代传统PC

2023年全球移动电话出货量约为**11.7亿部**，PC出货量约为**2.4亿台**



移动终端大规模普及催生多样化应用



新的算力需求

世界正在进入万物互联的时代



海量数据处理需要大带宽高并发

处理器：x86 -> ARM

部署模式：PC应用 -> 移动应用 -> 移动应用云化

移动终端接入方式：3G -> 4G -> 5G

需要云数据中心侧与端侧同构的算力

需要高性能、高并发、高吞吐的算力

AI算力占数据中心算力80%

边缘侧：数据采集 -> 数据实时智分析

云数据中心侧：承载海量数据的分析、处理和存储

云边协同：中心训练+边缘推理

当前计算产业呈现两个大的变化趋势：

- 移动智能终端取代传统PC：2023年全球传统PC出货量2.4亿台，而2023年全球移动智能终端出货量为11亿部。这导致计算正在从x86->ARM；应用正从PC应用->移动应用->移动应用云化。新的算力需求：云数据中心侧与端侧同构的算力。
- 世界正在进入万物互联的时代：2018年全球联接设备的数量已超过230亿；到2025年预计这一数字将突破1000亿，将带来海量的数据。例如：自动驾驶每天产生64 TB数据、深圳平安城市每天1,000 TB数据。

对于海量数据的处理的算力需求：

- 边缘侧实时智能处理，需要AI的算力。
- 数据中心侧分析、处理和存储海量的数据，需要高并发、高性能、高吞吐的算力。

计算产业的商业动力

- 超万亿规模的计算产业空间。
- 新应用、新技术、新计算架构，百亿级联接、爆炸式数据增长将重塑ICT产业新格局，催生新的计算产业链条，涌现出新的厂家和新的生态体系：
 - 软件：操作系统和虚拟化软件、数据库、中间件、大数据平台、企业应用软件，云服务、数据中心管理服务。
 - 硬件：服务器及部件、企业存储设备。
 - 云服务、数据中心管理服务。

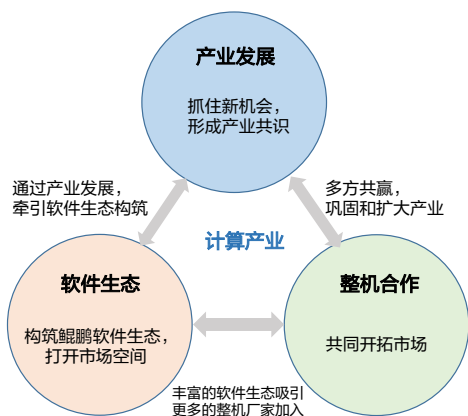


2023年全球计算产业投资额（美元）

- 数据来源：IDC

构筑软件生态，打开市场空间

愿景：驱动计算产业变革，达成鲲鹏计算产业共识，让世界有第二种选择



产业发展

影响国家/区域政策（影响政策方向，规避政策风险）
引领行业方向（参与行业及事实标准制定）

探索前沿科技，培育行业影响力
联合创新与开源，建设开源生态

软件生态

云与互联网软件：最优性价比驱动云与互联网产业生态汇聚

行业专用软件：运营商、金融、政府安平等行业国产化替代

基础平台软件：与OS、数据库、中间件软件厂家行程商业联盟

整机合作

OEM/ODM/集成商：通过技术、项目合作，提供技术支持，共同拓展

渠道合作伙伴：对下游经销商进行管理赋能、技术支持，共同扩展

思考题

1. （单选题）以下哪一项不是人工智能应用落地的要素？（ ）
 - A. 数据
 - B. 场景
 - C. 算力
 - D. 人才
2. （多选题）印加人使用结绳计数时，能够表示出以下哪些选项中的数字单位？（ ）
 - A. 百
 - B. 万
 - C. 十万
 - D. 千

- 参考答案：
- 1： D
- 2： ABD

本章总结

- 本章我们了解了计算的发展历程是怎么样的、对人类的计算是如何一步步发展至今天有了大致的了解，同时通过对计算产业的发展趋势的学习，我们了解了目前计算产业的几个主要发展方向，遇到的挑战和发展前景。
- 在后续的课程中我们将继续探索计算系统架构概述，逐步带领大家走进计算的世界，智能计算的介绍之旅正式开启。

学习推荐

- 鲲鹏社区：
 - <https://www.hikunpeng.com/>
- 昇腾社区：
 - <https://www.hiascend.com/zh/>

缩略语

- BMC: Baseboard Management Controller, 主板管理控制单元, 负责各路传感器的信号采集、处理、储存, 以及各种器件运行状态的监控。
- BMS: Bare Metal Server, 裸金属服务器, 提供单租户专属的物理服务器。
- CISC: Complex Instruction Set Computing, 复杂指令集计算机, 一种CPU指令集。
- DSP: Digital Signal Processing, 数字信号处理, 将模拟信号转化为二进制数字信号并进行处理。
- GPU: Graphic Processing Unit, 图形处理器, 执行复杂的数学和几何图形计算的计算单元。
- HPC: High-Performance Computing, 高性能计算, 专用于处理大型计算问题的计算机集群系统。
- CPU: Central Processing Unit, 中央处理器, 是计算机的通用计算单元。

缩略语

- NIST: National Institute of Standards and Technology, 美国国家标准与技术研究院, 美国商业部门技术管理组织内部的一个联邦代理机构。
- IDC: international data company, 国际数据公司, 一家信息技术领域咨询服务提供商。
- AI: artificial intelligence, 人工智能, 它是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学。
- CRM: Computing Resource Manager, 计算资源管理器, 主要用于管理集群内的计算节点所映射的虚拟计算资源。
- ERP: Enterprise Resource Planning, 企业资源计划, 一个建立在信息技术基础上, 为企业员工及决策层提供决策手段的管理平台。
- NFV: Network Functions Virtualization, 网络功能虚拟化, 以软件方式实现网络功能的技术。
- CANN: Compute Architecture for Neural Networks, 昇腾异构计算架构, 一种针对AI场景推出的异构计算架构。

Thank you.

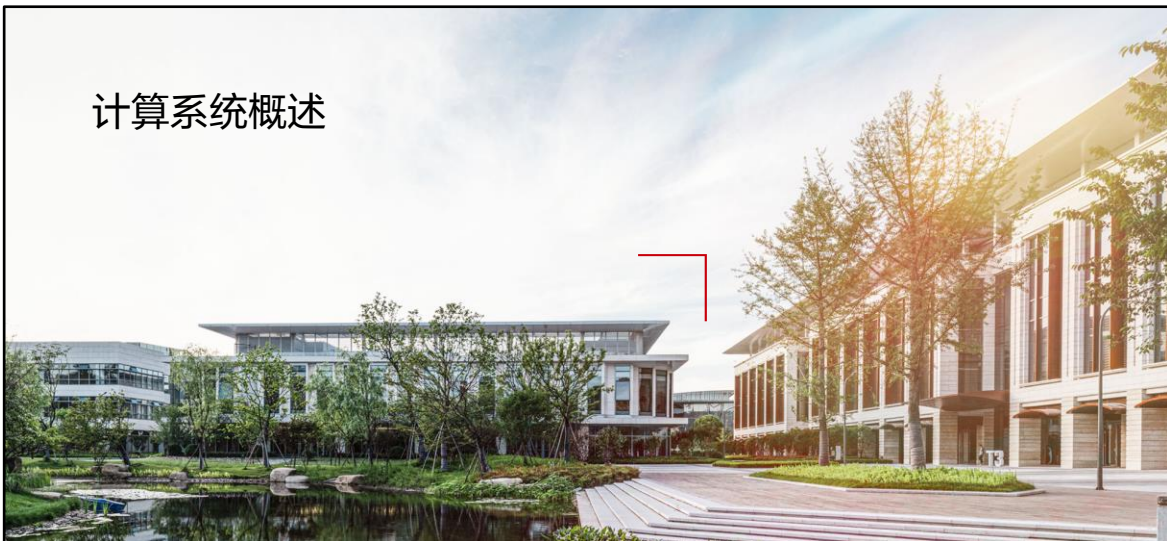
把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。
Bring digital to every person, home, and
organization for a fully connected,
intelligent world.

Copyright©2024 Huawei Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.



计算系统概述



前言

- 通过之前的学习，我们对计算产业的现状及发展有了大致的了解，那么在计算产业中，计算产品的服务器的定义是什么？服务器有哪些分类？上层又有哪些计算系统代表软件？在软硬件的支持下又实现了哪些解决方案呢？
- 本章我们将带着这些问题，一起去学习计算系统的架构、硬件、软件和解决方案。

目标

- 学完本课程后，您将能够：
 - 熟悉计算系统基本架构
 - 掌握计算产品分类
 - 掌握计算系统组成

目录

1. 计算系统架构
2. 计算系统硬件
3. 计算系统软件
4. 计算系统解决方案

计算系统简介

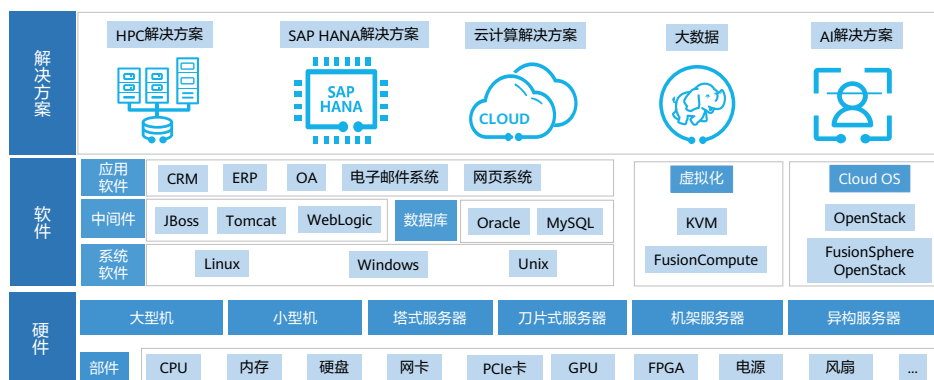
- 计算系统是一个复杂的系统，包括硬件、软件以及相关的服务和资源，按照用户的需求对数据进行存储、处理、计算并给出输出结果。
- 随着计算系统的不断发展，计算系统的能力也变得更强大，结构也变得更复杂。例如各种先进的管理、加速、异构、AI等系列芯片，现已被融入覆盖云、边、端的全栈全场景智能解决方案的计算系统体系架构中了。
- 可以看出，计算系统已经成为了计算业务全方位智能化加速的根本与基石。



Computation System = Computer + Data

- 狭义上的计算系统由硬件系统和软件系统构成。硬件系统包括主机和外设设备，软件系统包括系统软件和应用软件。

计算系统架构全景图概览



- 计算架构全景图概览如图所示，由于计算系统架构庞大且复杂，因此本页中将计算系统大致划分为硬件、软件还有解决方案这三个层次。
- 本章的后续的内容将会着重围绕着本图中的三个层次内的重要部分进行简要介绍，帮助大家理解什么是计算系统硬件，比如大型机，小型机还有其他的服务器种类是如何分类的。之后再通过对计算系统软件的学习，了解什么是服务器系统软件，中间件，虚拟化等等。在对计算系统软硬件有了初步了解之后，本章后续会对一些常见的解决方案进行介绍，让学员了解在有了软硬件之后，我们能够做出什么样的解决方案来满足哪些场景的需求，以满足人们现实中工作生活的需要。

目录

1. 计算系统架构
- 2. 计算系统硬件**
3. 计算系统软件
4. 计算系统解决方案

计算机系统硬件——服务器

- 服务器定义

- 服务器是计算机的一种。它比普通计算机运行速度更快、负载更高且价格更贵。
- 服务器是为用户提供服务的计算机，通常分为文件服务器、数据库服务器和应用程序服务器等。



- 服务器是20世纪90年代迅速发展主流计算产品，能为网络用户提供集中计算、信息发布及数据管理等服务，也可以将与其相连的如硬盘、打印机、Modem等各种专用通讯设备给网络上的用户进行共享。
- 服务器的主要特点包括：
 - R: Reliability—可靠性，能连续正常运行多长时间。
 - A: Availability—可用性，系统正常运行时间和使用时间的百分比。
 - S: Scalability—可扩展性，包括两方面，一方面是硬件的可扩展性，另一方面是软件对操作系统的支持能力。
 - U: Usability—易用性，服务器的硬件和软件易于维护和修复。
 - M: Manageability—可管理性，对服务器运行情况能进行监控，报警，对一些故障的自动智能化处理。

服务器发展历程

大型、小型机阶段



微型计算机阶段



通用服务器阶段

- 大型主机阶段：
20世纪40-50年代，ENIAC的诞生开启了大型主机的时代。
- 小型计算机阶段：
20世纪60-70年代，是对大型主机进行的第一次“缩小化”，可以满足中小企业事业单位的信息处理要求，成本较低，价格可被接受。
- 微型计算机阶段：
20世纪70-80年代，是对大型主机进行的第二次“缩小化”，1976年美国苹果公司成立，1977年推出了Apple II计算机。1981年IBM推出IBM-PC。
- 通用服务器时代：
20世纪80年代，康柏推出全球第一款x86架构标准服务器：SystemPro。
20世纪90年代，英特尔推出Pentium Pro—为服务器而生的x86处理器。

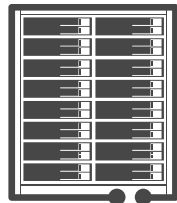
- 大型主机阶段：
 - 20世纪40-50年代，ENIAC的被研发出来，它长30米多，宽6米，高约2米多，重30多吨，它的诞生开启了大型主机的时代。随着时代的发展，现在的大型机的体积已经小很多了，它们大多高约2米，长约1到3米，深约1米。
- 小型计算机阶段：
 - 20世纪60-70年代，是对大型主机进行的第一次“缩小化”，可以满足中小企业事业单位的信息处理要求，成本较低，价格可被接受。小型机的体积虽然比大型机小了不少，但大多数也还有个立柜那么大。
- 微型计算机阶段：
 - 20世纪70-80年代，是对大型主机进行的第二次“缩小化”，这个时代的个人计算机的体积已经缩小到可以放在人们的办公桌面上了，1976年美国苹果公司成立，1977年就推出了Apple II计算机，大获成功。1981年IBM推出IBM-PC，此后它经历了若干代的演进，占领了个人计算机市场，使得个人计算机得到了很大的普及。
- 通用服务器时代：
 - 1978年，英特尔推出第一代x86架构处理器。1993年，英特尔正式推出Pentium（奔腾）系列，该系列的推出，将x86架构处理器带上了一个新的性能高度。
 - 1995年，英特尔推出Pentium Pro—为服务器而生的x86处理器，从此开启了x86的至强时代，其标准化开放性也促成了市场发展，为云计算时代打下了坚实物质基础。
- 经历了电子管数字计算机、晶体管数字计算机、集成电路数字计算机和大规模集成电路数字计算机的发展历程，计算机技术逐渐走向成熟。

服务器分类：按形态划分

塔式服务器



刀片式服务器



机架式服务器



- 塔式服务器：
 - 立式放置的服务器机型，外形以及结构都跟立式PC差不多。
 - 主板扩展性较强，插槽数多，因而应用范围广泛，可以满足一般常见的服务器应用需求。
- 刀片式服务器：
 - 刀片式服务器是指在标准高度的机架式机箱内可插装多个卡式的服务器单元，实现高可用和高密度。
 - 每一块“刀片”实际上就是一个服务器。
- 机架式服务器：
 - 机架服务器是一种外观按照统一标准设计的服务器，配合机柜统一使用。
 - 是一种优化结构的塔式服务器。

塔式服务器

- 塔式服务器：
 - 立式放置的服务器机型，外形以及结构都跟立式PC差不多。
 - 主板扩展性较强，插槽数多，因而应用范围广泛，可以满足一般常见的服务器应用需求。

优点

- 便于维护升级：塔式服务器的主板通常会预留扩展插槽，机箱内部也会预留空间来方便后续硬盘，内存等元件的扩展，更换或升级。
- 成本较低：塔式服务器的成本较低，中小企业或个体用户也能负担。

缺点

- 占用空间大：由于单台塔式服务器体积较大，在需要多台服务器同时工作以支撑高负载应用时，需要较多的空间，管理起来也不方便。
- 较高能耗：塔式服务器独立的供电和散热系统在带来了良好的散热效果的同时，也使得它的能耗比机架式服务器的能耗更高。

刀片式服务器

- 刀片式服务器：

- 刀片式服务器是指在标准高度的机架式机箱内可插装多个卡式的服务器单元，实现高可用和高密度，它是专门为特殊应用行业和高密度计算环境设计的。
- 每一块“刀片”就是一个个独立的服务器，每块刀片母板都运行自己的操作系统，相互独立，可以服务不同的用户。如果是应对高负载的场景，母板们可以被系统软件整合为一个集群，在集群模式下，所有的母板可以连接起来提供高速的网络环境，可以共享资源，为相同的用户群服务。

优点

- 占用空间小：多个刀片母板被高度集成在一个机箱中，他们共享基本组件和电源，便于在数据中心中部署和管理。
- 易于扩展：刀片式服务器支持热插拔，可以根据用户的需要灵活的添加或移除刀片服务器，用户可以通过统一的管理平台来对这些服务器进行集中管理和监控。

缺点

- 散热要求高：由于刀片式服务器的紧凑性，刀片母板排列在一起很容易积热，过热可能会导致服务器受损。
- 成本高：刀片服务器的价格高，安装维护也会更加复杂，这使得它的部署和维护成本都较高。

机架式服务器

- 机架式服务器

- 机架服务器是一种外观按照统一标准设计的服务器，配合机柜统一使用。
- 机架式服务器有1U（1U=4.45cm）、2U、4U等规格。机架式服务器通常安装在标准的19英寸机柜里面，被部署在中大型企业的数据中心中，以应对像互联网和金融行业等需要部署大量服务器来提供服务的场景。

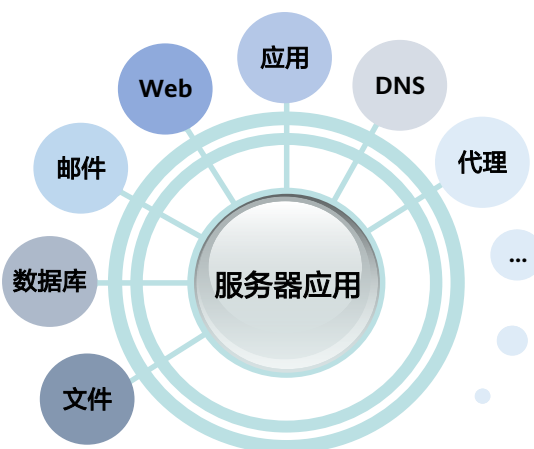
优点

- 节省空间：机架式服务器可以整齐有序的被安装在数据中心的机柜中，更易于集中管理。
- 能耗较低：多台机架服务器可以共享同一套散热系统，能效比高。

缺点

- 扩展能力有限：相比塔式服务器，机架式服务器的扩展空间更小，部件可能不容易升级和维护。
- 初始成本高：建设用于部署机架式服务器的机柜，电力系统，冷却系统需要更多的投资。

服务器分类：按用途划分



- 服务器已经广泛应用在电信运营商、政府、金融、教育、企业、电子商务等各个行业领域，为用户提供文件、数据库、邮件、Web等服务。
- 服务器应用部署架构：
 - C/S：Client/Server的缩写。通常也称为客户端/服务器架构。服务器端运行服务端程序，客户端安装客户端软件。在此架构里服务端和客户端分别完成不同的任务，客户端处理用户的前端界面和交互操作，服务端处理后台业务逻辑和请求数据，这使得两端的通讯速度和通讯的效率大大提高。例如，我们在文件服务器上（服务器端）安装vsftpd程序，并启动服务；在用户的计算机中安装FileZilla或WinSCP等客户端工具后，用户就可以通过客户端工具进行文件的上传和下载。
 - B/S：Browser/Server的缩写。通常也称为浏览器/服务器架构。B/S架构中，用户只需安装浏览器即可，而将应用逻辑集中在服务器和中间件上，可以提高数据处理性能。例如，我们在访问一个网站时，只需要在自己的浏览器中输入网站的域名，如www.huawei.com，就可以看到该网站的后台服务器给我们提供的Web服务。而网站的后台服务器端可能有很多服务器提供服务，如数据库服务，代理服务，缓存服务等等，这些都不需要用户关心，用户只需要通过浏览器输入网址就可以看到相应的界面。

服务器分类：按处理器类型

- 目前主流服务器按照搭载的处理器类型，可被分为x86架构服务器和ARM架构服务器，x86架构处理器基于CISC复杂指令集，而ARM架构处理器基于RISC精简指令集。除此之外较为常见的处理器架构还有RISC-V和MIPS。



指令集类型	处理器架构	应用场景	产品示例
CISC	x86	常用于个人电脑及服务器芯片	Intel Core CPU，AMD RYZEN CPU
RISC	ARM	常用于服务器，移动和嵌入式设备中	华为鲲鹏处理器，苹果A系列处理器
RISC	RISC-V	常用于可穿戴设备和嵌入式MCU	高通Snapdragon Wear芯片
RISC	MIPS	常用于住宅网关和路由器设备中	龙芯2处理器

处理器指令集：CISC架构与RISC架构

- CISC (Complex Instruction Set Computer, 复杂指令集计算机)：早期的CPU全部是CISC架构，它的设计目的是要用最少的机器语言指令来完成所需的计算任务。指令数目多而复杂，执行每条指令的时间长，但单条指令可以处理的工作较为丰富。这种架构会增加CPU结构的复杂性和对CPU工艺的要求，但对于编译器的开发十分有利。
- RISC (Reduced Instruction Set Computer, 精简指令集计算机)：RISC架构要求软件来指定各个操作步骤。指令集较为精简，单个指令的运行短，执行操作简单又高效，但复杂命令需要使用多条指令实现，这种架构可以降低CPU的复杂性以及允许在同样的工艺水平下生产出功能更强大的CPU，但对于编译器的设计有更高的要求。
- 基于CISC架构开发的软件不能直接运行在RISC架构的平台上，反之亦然。不同指令集的指令格式的和编码方式不同，导致相同的程序无法直接在不同的指令集上运行。

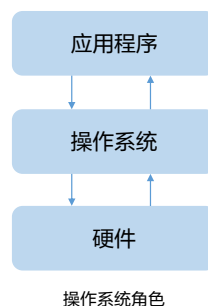
- 随着VLSI (超大规模集成电路) 技术的发展，为了让软件编程更加方便高效，提高程序的运行速度，工程师不断增加可实现复杂功能的指令和多种灵活的编址方式。为了实现复杂操作，微处理器除了提供各种寄存器和机器指令功能之外，还通过存储在只读存储器 (ROM) 中的微程序来处理并分析每一条指令，之后执行一系列的初级指令来实现要求的功能，这种设计的形式架构被称为CISC结构。
- 随着要实现的功能越来越多，构建一个日趋繁杂的指令系统难以实现，甚至有可能降低系统性能。在1979年以帕特逊教授为首的一批科学家在美国加州大学伯克莱分校开展研究时发现，CISC中各种指令的使用率相差甚远，例如一个典型的运算过程所使用的80%的指令，只占到指令系统的20%，最被频繁被用到的指令其实是取、存、和加这样的简单指令，这就导致复杂指令系统的设计浪费。针对这个问题，帕特逊教授提出了设计一个精简指令架构的设想，只包含那些使用频率很高的少量指令，再加上一些必要的指令去支持操作系统和高级语言，遵循着这个设计理念而发展而来的指令集架构就是RISC。

目录

1. 计算系统架构
2. 计算系统硬件
- 3. 计算系统软件**
4. 计算系统解决方案

计算系统软件——操作系统

- 计算机主机由各种硬件所组成，为了有效的控制这些硬件资源，人们开发了操作系统，目前主流服务器操作系统主要基于Linux，Windows。
- 操作系统本质上是管理计算机硬件与活动的系统软件，重点负责管理驱动系统中的各种硬件设备。操作系统可以获取上层应用程序发来的处理需求，并传递给下层硬件，让CPU开始执行逻辑判断和数学运算，让内存开始写入和读取数据或程序，让硬盘开始存储数据，让各种各样的硬件开始工作。



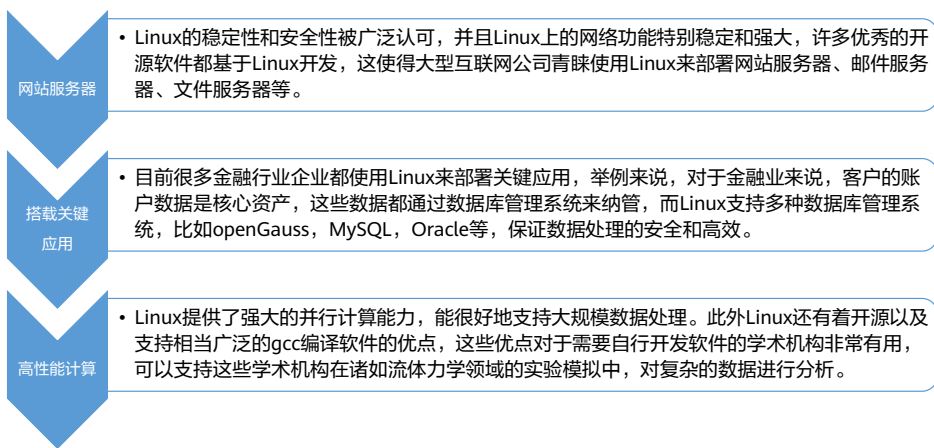
- 操作系统主要由系统调用和内核两部分组成，内核主要用于管理硬件与提供对应的功能上（例如控制CPU运行，内存读写，硬盘读写等等），由于内核十分重要，如果用户能够直接使用到内核的话，很容易会误操作损坏内核从而导致系统崩溃，为了保护内核也为了让工程师们比较容易开发软件，操作系统会提供一套API，也就是系统调用层，工程师们只要遵循系统调用提供的接口和参数来开发软件，软件就可以在内核上运行。

Linux

- Linux是一个遵循GNU GPL授权模式的开源操作系统，它包含操作系统最底层的内核及其提供的内核工具，任何人均可取得源代码并执行或修改它的内核程序，不同的商业公司或非盈利性质的团体将Linux内核及其工具整合起来，再加上自身独有的一些工具程序，对外发布，就有了Linux发行版。常见的Linux发行版有openEuler、Red Hat、SUSE、Ubuntu等。
- Linux有着开源免费、性能稳定、内核小、支持多用户多任务以及使用容易等优点，并且还在不断迭代更新，这使得它成为了最为流行的服务器操作系统之一，同时也被广泛的应用在软件开发及嵌入式系统中。



Linux服务器的应用场景

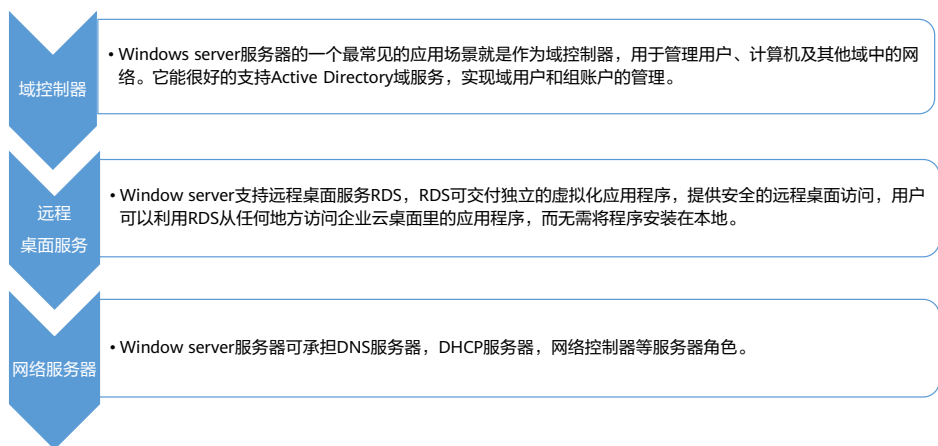


Windows Server

- Windows Server是由微软公司开发的服务器操作系统，用于构建连接的应用程序、网络 and Web服务的基础结构，包括从工作组到数据中心。
- Windows Server目前有两个主要的发布渠道：长期服务频道（LTSC）和年度渠道（AC），长期服务频道注重对于操作系统的基本功能和系统安全进行维护更新，而年度渠道通过提升发布频率，使得用户能够更快的用上一些创新功能，例如会更关注像容器和微服务这样的新功能的迭代更新。
- 目前Windows Server 2022主要提供标准版（Standard）和数据中心版（Datacenter）。

- 以Windows Server 2022为例，数据中心版和标准版主要的差异有：相较于标准版，数据中心版增加了对软件定义网络 and Storage Spaces Direct的支持，每个许可证可允许运行无限台虚拟机以及一台Hyper-V主机（标准版每个许可证仅能允许运行2台虚拟机以及一台Hyper-V主机），并能够作为网络控制器了。

Windows Server服务器的应用场景



计算系统软件——中间件

- 中间件（Middleware），是一类提供系统软件和应用软件之间连接、便于软件各部件之间的沟通的软件，应用软件可以借助中间件在不同的技术架构之间共享信息与资源。中间件位于客户机服务器的操作系统之上，管理着计算资源和网络通信。
- 中间件能为上层应用的开发提供即开即用的服务交互与计算的能力，提高开发效率并大幅缩短开发周期，中间件通过屏蔽底层runtime的差异，使得应用程序无需感知底层的复杂性，节省系统资源，减少运行成本。

——维基百科

- Runtime指为了支持程序运行所需要提供的运行时库（runtime library），以及运行环境（Runtime environment），又称运行时系统（run-time system）每种语言的Runtime都不一样，例如常见的运行环境有JRE，C#，Visual Basic .NET等等。不同高级语言的Runtime的功能和实现方法有差异。

中间件概述

JBoss

- JBoss是一个基于Java的开源应用服务器，同时它也是一套开源的企业级Java中间件系统，它提供了一套工具和服务，用于实现基于SOA的Java企业应用和服务。
- 由于JBoss遵循LGPL许可，其他公司可以在任何商业应用中免费使用，这使得JBoss应用服务器发展十分迅猛。

Tomcat

- Tomcat是一个由Apache软件基金会、Sun、其他公司以及个人共同开发而成的免费的、开放源代码的Web应用服务器。
- Tomcat是一个小型的轻量级应用服务器，支持最新的Servlet和JavaServer Page (JSP) 规范，并且技术先进、性能稳定、扩展性好、运行时占用的系统资源小，因此在中小型系统和并发访问用户不多时被普遍应用。

WebLogic

- WebLogic是由Oracle公司出品的Java应用服务器，与JBoss类似，它也是基于JEE的中间件。
- WebLogic的功能是对适用于本地环境和云环境的企业Java应用进行开发、部署和运行。

- JBoss遵循JEE（Java Platform, Enterprise Edition），支持EJB（Enterprise JavaBeans）1.1、EJB2.0和EJB3的规范。
- JBoss和Tomcat是开源免费的，WebLogic与另一个IBM集成软件平台WebSphere则是闭源收费的。
- Tomcat作为小型的轻量应用服务器，在中小型和并发访问用户不多时被普遍应用，但Tomcat不支持EJB，对EJB有需求的中小型公司可以选用JBoss。
- WebLogic和WebSphere作为商业软件有着功能强大，高扩展性的特点，主要应用于大型企业项目。

计算系统软件——数据库

- 数据库是指长期存储在计算机内、有组织、可共享的大量的数据的集合。数据库中的数据按照一定的数据模型组织、描述和存储，具有较小的冗余度、较高的数据独立性和易扩展性，并可为各种用户共享。



永久存储

有组织

可共享

- 数据库（Database）指长期存储在计算机内的、有组织的、可共享的数据集合。通俗的讲，数据库就是存储数据的地方，就像冰箱是存储食物的地方一样。在生活中，每个人都在使用数据库。当我们在电话簿里查找名字时，就是在使用数据库。在某个浏览器上进行搜索时，也是在使用数据库。以及平时我们登录网络，也需要依靠数据库验证自己的名字和密码。即便是在使用 ATM 机时，也要利用数据库进行 PIN 码验证和余额检查。
- 数据库实际上就是一个文件集合，是一个存储数据的仓库，本质就是一个文件系统，数据库是按照特定的格式把数据存储起来，用户可以对存储的数据进行增删改查操作。
- 在日常生活中，人们可以直接用中文、英文等自然语言描述客观事物。在计算机中，则要抽象出对这些事物感兴趣的特征，并组成一个记录来描述。
- 数据库是结构化信息或数据（一般以电子形式存储在计算机系统中）的有组织的集合。
- 数据库，简而言之可视为电子化的文件柜——存储电子文件的处所，用户可以对文件中的数据运行新增、截取、更新、删除等操作。
- 所谓“数据库”是以一定方式储存在一起、能予多个用户共享、具有尽可能小的冗余度、与应用程序彼此独立的数据集合。

为什么要使用数据库？

- 使用数据库可以高效且条理分明地存储数据，它使人们能够更加迅速和方便地管理数据，主要体现在以下几个方面。

1. 数据库可结构化存储大量的数据信息，方便用户进行有效的检索和访问。
2. 数据库可有效地保持数据信息的一致性、完整性、降低数据冗余。
3. 数据库可满足应用的共享和安全方面的要求，把数据放在数据库中在很多情况下也是出于安全的考虑。
4. 数据库技术能够方便智能化地分析，产生新的有用信息。



- 随着互联网技术的高速发展，预计2020年底全世界网民的数量将达到50亿。网民数量的增加带动了网上购物、微博，网络视频等产业的发展。那么，随之而来的就是庞大的网络数据量。大量的数据正在不断产生，那么如何安全有效地存储、检索，管理它们呢？于是对数据的有效存储、高效访问、方便共享和安全控制等问题成为了信息时代一个非常重要的问题。
- 数据库可以结构化存储大量的数据信息，方便用户进行有效的检索和访问：
 - 数据库可以对数据进行分类保存，并且能够提供快速的查询。例如，我们平时使用百度搜索内容时，百度也是基于数据库和数据分类技术来达到快速搜索的目的。
- 数据库可以有效地保持数据信息的一致性、完整性、降低数据冗余：
 - 可以很好地保证数据有效、不被破坏，而且数据库自身有避免重复数据的功能，以此来降低数据的冗余。
- 数据库可以满足应用的共享和安全方面的要求，把数据放在数据库中在很多情况下也是出于安全的考虑：
 - 例如，如果把所有员工信息和工资数据都放在磁盘文件上，则工资的保密性就无从谈起。如果把员工信息和工资数据放在数据库中，就可以只允许查询和修改员工信息，而工资信息只允许指定人（如财务人员）查看，从而保证数据的安全性。
- 数据库技术能够方便智能化地分析，产生新的有用信息：
 - 例如，超市中把物品销售信息保存在数据库中，每个月销售情况的排名决定了下半月的进货数量。数据库查询的结果实际上产生了新的数据信息。

数据库管理系统 DBMS

- 数据库管理系统（DataBase-Management System，DBMS）由一个互相关联的数据的集合和一组用以访问这些数据的程序组成。这个数据集合通常被称作数据库（Database）。DBMS的主要目的是提供一种可以方便、高效地存取数据库信息的途径。
- DBMS充当数据库与其用户或程序之间的接口，允许用户检索、更新和管理信息的组织和优化方式。此外，DBMS还有助于监督和控制数据库，提供各种管理操作，例如性能监视、调优、备份和恢复。
- 具有代表性的数据库管理系统有：Oracle、Microsoft SQL Server、Access、MySQL及PostgreSQL等。

- 数据库通常离不开完备的数据库软件程序，也就是数据库管理系统（DBMS）DBMS充当数据库与其用户或程序之间的接口，允许用户检索、更新和管理信息的组织和优化方式。此外，DBMS还有助于监督和控制数据库，提供各种管理操作，例如性能监视、调优、备份和恢复。
- 数据库软件旨用于创建、编辑和维护数据库文件及记录，帮助用户更轻松地执行文件和记录创建、数据录入、数据编辑、更新和报告等操作。除此之外，数据库软件还能处理数据存储、备份和报告以及多路访问控制和安全性等问题。随着当今数据盗窃日益频繁，数据库安全性已变得至为重要。数据库软件有时也称为“数据库管理系统”（DBMS）。

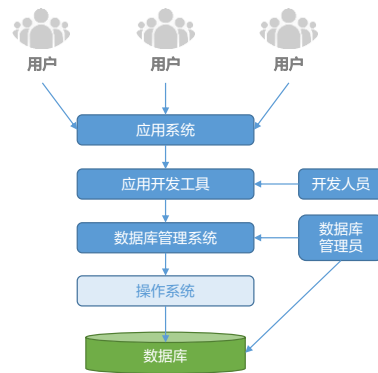
数据库管理系统 – 功能介绍

- **数据定义功能：**DBMS提供数据定义语言（Data Definition Language，DDL），用户通过它可以方便地对数据库中的数据对象进行定义。
- **数据操纵功能：**DBMS还提供数据操纵语言（Data Manipulation Language，DML），用户可以使用DML操作数据，实现对数据库的基本操作，如查询、插入、删除和修改等。
- **数据库的运行管理：**数据库在建立、运用和维护时由数据库管理系统统一管理、统一控制，以保证数据的安全性、完整性、多用户对数据的并发使用及发生故障后的系统恢复。
- **提供方便、有效地存取数据库信息的接口和工具：**编程人员可通过编程语言与数据库之间的接口进行数据库应用程序的开发。数据库管理员（Database Administrator，DBA）可通过提供的工具对数据库进行管理。
- **数据库的建立和维护功能：**数据库功能包括数据库初始数据的输入、转换功能，数据库的转储、恢复功能，数据库的重组功能和性能监控、分析功能等。这些功能通常是由一些程序来完成。

- **数据库的运行管理：**数据库在建立、运用和维护时由数据库管理系统统一管理、统一控制，以保证数据的安全性、完整性、多用户对数据的并发使用及发生故障后的系统恢复。例如：
 - 数据的完整性检查功能保证用户输入的数据应满足相应的约束条件。
 - 数据库的安全保护功能保证只有赋予权限的用户才能访问数据库中的数据。
 - 数据库的并发控制功能使多个用户可以在同一时刻并发地访问数据库的数据。
 - 数据库系统的故障恢复功能使数据库运行出现故障时可以进行数据库恢复，以保证数据库可靠地运行。

数据库系统 DBS

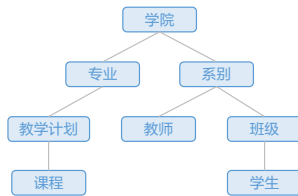
- 数据库系统（Database System，DBS）由硬件和软件共同构成。硬件主要用于存储数据库中的数据，包括计算机、存储设备等。软件部分主要包括数据库管理系统、支持数据库管理系统运行的操作系统，以及支持多种语言进行应用开发的访问技术等。
- 数据库系统是指在计算机系统中引入数据库后的系统。一个完整的数据库系统一般由数据库、数据库管理系统、应用开发工具、应用系统、数据库管理员和用户组成。



- 数据库系统主要有以下 3 个组成部分：
 - 数据库：用于存储数据的地方。
 - 数据库管理系统：用于管理数据库的软件。
 - 数据库应用程序：为了提高数据库系统的处理能力所使用的管理数据库的软件补充。
- 数据库（DataBase，DB）提供了一个存储空间来存储各种数据，可以将数据库视为一个存储数据的容器。一个数据库可能包含许多文件，一个数据库系统中通常包含许多数据库。
- 数据库管理系统（Database Management System，DBMS）是用户创建、管理和维护数据库时所使用的软件，位于用户和操作系统之间，对数据库进行统一管理。DBMS 能定义数据存储结构，提供数据的操作机制，维护数据库的安全性、完整性和可靠性。虽然已经有了 DBMS，但是在很多情况下，DBMS 无法满足对数据管理的要求。
- 数据库应用程序（DataBase Application）的使用可以满足对数据管理的更高要求，还可以使数据管理过程更加直观和友好。数据库应用程序负责与 DBMS 进行通信、访问和管理 DBMS 中存储的数据，允许用户插入、修改、删除数据库中的数据。

数据库模型 – 层次、网状、关系模型

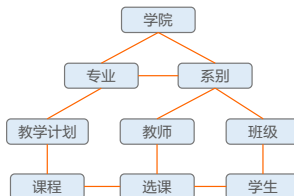
层次数据模型



- 层次模型

- 有且只有一个节点没有双亲，该节点被称为根节点（root）。
- 根节点以外的其他节点有且只有一个双亲节点。

网状数据模型



- 网状模型

- 允许一个以上的节点无双亲。
- 一个节点可以有多个的双亲。

关系数据模型



- 关系模型

- 建立在严格的数据概念基础上。
- 关系必须是规范化的。
- 关系的分量必须是一个不可分的数据项。

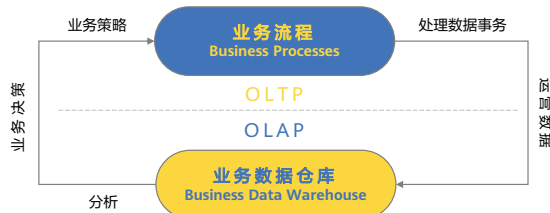
- 层次数据库（Hierarchical Database, HDB）：

- 层次数据库是最早研制成功的数据库系统，它把数据通过层次结构（树形结构）的方式表现出来。层次数据库曾经是数据库的主流，但随着关系数据库的出现和普及，现在已经很少使用了。比较具有代表性的层次数据库是IMS（Information Management System）数据库，由IBM公司研制成功。

- 关系型数据库（Relational Database, RDB）：

- 关系型数据库是现在应用最广泛的数据库。关系型数据库在1969年诞生，可谓历史悠久。和Excel工作表一样，关系型数据库也采用由行和列组成的二维表来管理数据，所以简单易懂。同时，它还使用SQL（Structured Query Language，结构化查询语言）对数据进行操作。传统的关系型数据库采用表格的存储方式，数据以行和列的方式进行存储，要读取和查询都十分方便。

关系型数据库主流应用场景



联机事务处理 OLTP On-Line Transaction Processing

- OLTP是传统的关系型数据库的主要应用，事务性非常高的系统，一般都是高可用的在线系统，以小的事务以及小的查询为主，主要是基本的、日常的事务处理，例如银行交易。
- OLTP 系统强调数据库内存效率，评估其系统的时候，一般看其每秒执行的Transaction以及Execute SQL的数量。强调内存各种指标的命中率，强调绑定变量，强调并发操作。

联机分析处理 OLAP On-Line Analytical Processing

- OLAP是数据仓库系统的主要应用，支持复杂的分析操作，侧重决策支持，并且提供直观易懂的查询结果。在这样的系统中，语句的执行量不是考核标准，因为一条语句的执行时间可能会非常长，读取的数据也非常多。
- OLAP 系统则强调数据分析，考核的标准往往是磁盘子系统的吞吐量（带宽），强调SQL执行时长，强调磁盘I/O，强调分区等。

- OLTP是传统关系数据库的主要应用：面向基本的，日常的事务处理，例如银行储蓄业务的存取交易，转账交易等。
 - 特点：（1）大吞吐量：大量的短在线事务（插入、更新、删除），非常快速的查询处理；（2）高并发，（准）实时响应。
 - 典型的OLTP场景：零售系统；金融交易系统；火车机票销售系统；秒杀活动
- OLAP：联机分析处理的概念最早是E.F.Codd于1993年相对于OLTP系统而提出的。是指对数据的查询和分析操作，通常对大量的历史数据查询和分析。涉及到的历史周期比较长，数据量大，在不同层级上的汇总，聚合操作使得事务处理操作比较复杂。
 - 特点：（1）主要面向侧重于复杂查询，回答一些“战略性”的问题。（2）数据处理方面聚焦于数据的聚合，汇总，分组计算，窗口计算等“分析型”数据加工和操作。（3）从多维度去使用和分析数据。
 - 典型的OLAP场景：报表系统，CRM系统；金融风险预测预警系统、反洗钱系统；数据集市，数据仓库。

OLTP/OLAP 对比分析

特性	OLTP	OLAP
特征	操作处理	信息处理
面向	事务	分析
用户	办事员、DBA、数据库专业人员	分析决策人员
功能	日常操作	长期信息需求，决策支持
数据	当前的、确保最新	历史的，跨时间维护
汇总	原始的，高度详细	汇总的，统一的
工作单位	短的、简单事务	复杂查询
存取	读/写	大多为读
关注	数据进入	信息输出
操作	主关键字上索引/散列	大量扫描
访问记录数量	数十个	数百万
DB规模	100MB到GB	100GB到TB
优先	高性能，高可用性	高灵活性，端点用户自治
度量	事务吞吐量	查询吞吐量，响应时间

数据库对服务器的要求

- 数据库系统对于服务器的性能要求很高，加之数据库系统往往是企业IT系统里的核心，对于银行，证券等金融行业的IT系统来说就更是如此，数据库服务器必须保证在长期运作中维持高效运转，同时也要保护数据不丢失。服务器作为数据库的载体，需要提供给数据库系统必要的计算、网络还有存储等资源。
- 在选择服务器规格时，需要考虑以下一些关键指标：
 - 处理器：数据库处理数据时的负载并不是一成不变的，随着时间的变化必定会出现间隔性的出现负载高峰，选择服务器CPU时应该结合处理器架构特点，考虑数据库高峰期负载，选择规格较高的处理器来保证CPU的峰值负载最好始终不超过85%，让数据库能够快速响应。
 - 内存：由于数据库系统时时刻刻都要对大量数据进行读写，它的性能很大程度上取决于内存的大小和运行速率，因此应该选择足够的内存容量来满足数据库对大批量数据的读写需求。
 - 存储空间：在数据库中有大量的数据在处理完之后需要被保存，对于需要被频繁读取的数据来说需要将它们放置在一个空间足够大的本地硬盘或者是专门的存储设备中，方便快速取用。
 - 网络：数据库处理的大量数据大都是通过网络传输，为数据库配置一个乃至多个性能强劲的，能支持高网络带宽的网卡显得十分重要。

- 以openGauss数据库部署在4路鲲鹏服务器，通过软硬件协同优化配置达到openGauss数据库的极致性能的方法的实际场景为例，所需服务器规格如下：
 - 硬件规格如下：
 - 服务器： TaiShan 200(Model 2480)。
 - CPU： Kunpeng-920 ARM aarch64(4 Sockets * 64 Cores)。
 - 内存： 1TB。
 - 网卡： 万兆网卡Hi1822 Family(4*25GE)，时延 < 0.1ms。
 - 磁盘： NVME * 4， Model Number： HWE56P433T2M005N（V5 NVME卡）、HWE36P43016M000N（V3 NVME卡），其中V5 NVME卡1MB顺序写带宽达到2600MB以上。
 - 软件规格如下：
 - 操作系统： openEuler 20.03 (LTS)。
 - 数据库软件： openGauss 5.0.0 或其他更高Release版。

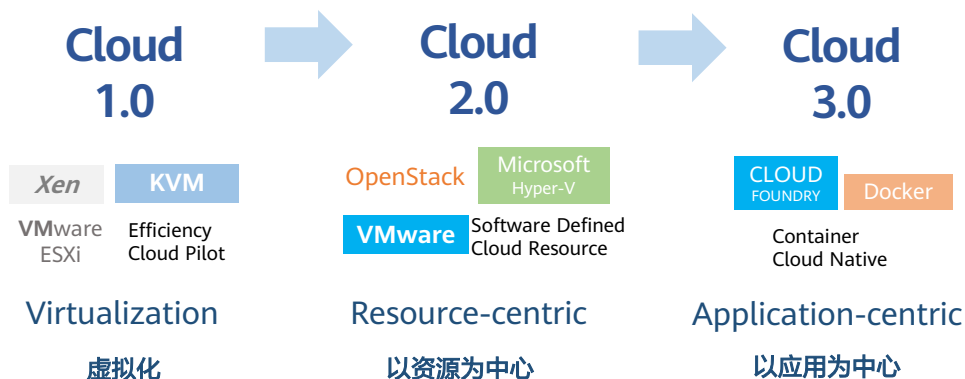
计算系统软件——Cloud OS

- 云操作系统（Cloud OS）是用于将物理设备以共享、弹性和可伸缩的方式进行供应和管理，提供人机界面，以可计量按需自助服务的方式为用户提供计算、存储和网络等基础云服务资源，可支撑PaaS、SaaS等服务执行的一种云计算系统软件。



云操作系统的发展

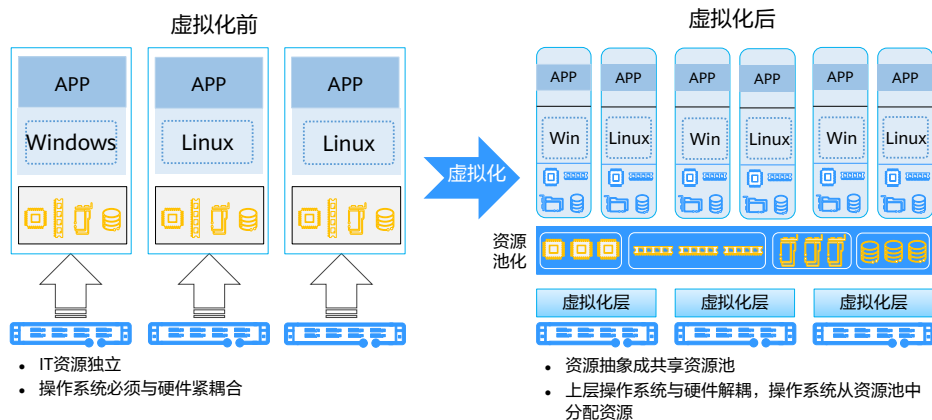
- 云操作系统的迭代发展可以归纳为由1.0版本的虚拟化，到2.0版本的以资源为中心的服务化，再到3.0的以应用为中心的云原生化。



- 针对上述三大云计算发展演进里程碑阶段而言，云计算1.0普遍已经是过去式。且一部分行业、企业客户已完成初步规模的云计算2.0建设商用，正在考虑该阶段的进一步扩容，以及面向云计算3.0的演进；而另一部分客户则正在从云计算1.0走向云计算2.0，甚至同步展开云计算2.0和3.0的演进评估与实施。
- Cloud 1.0：强调虚拟化，追求更高的资源利用率。
- Cloud 2.0：以资源为中心，实现基础设施云化、资源服务标准化、自动化。
- Cloud 3.0：以应用为中心，实现应用云化，敏捷应用开发与生命周期管理。

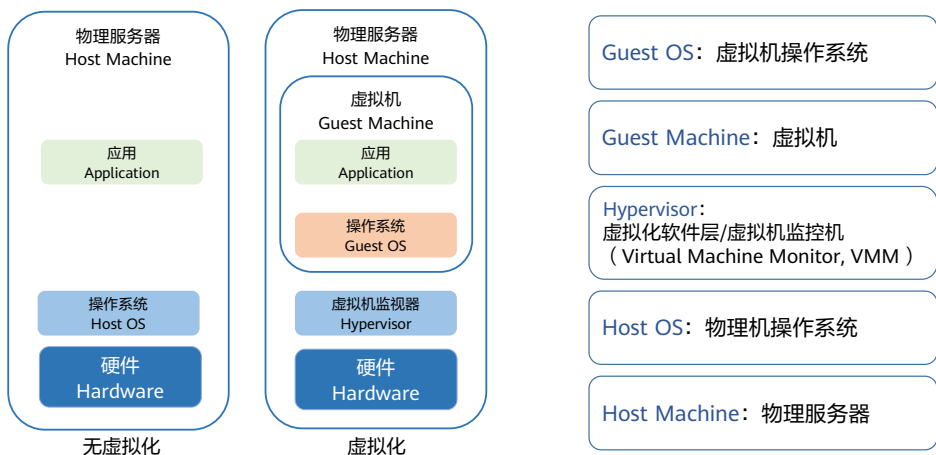
什么是虚拟化

- 虚拟化（Virtualization）的含义很广泛。将任何一种形式的资源抽象成另一种形式的技术都是虚拟化，是资源的一种逻辑表示。解除了物理硬件和操作系统之间的紧耦合关系。



- 虚拟化是云计算的基础。简单地说，虚拟化使得在一台物理的服务器上可以跑多台虚拟机，虚拟机共享物理机的CPU、内存、IO硬件资源，但逻辑上虚拟机之间是相互隔离的。
- 在计算机方面，虚拟化一般指通过对计算机物理资源的抽象，提供一个或多个操作环境，实现资源的模拟、隔离或共享等。
- 本质上，虚拟化就是对硬件资源的一种抽象与模拟。通过空间上的分割、时间上的分时以及模拟，虚拟化可将一份资源抽象成多份，亦可将多份资源抽象成一份。

虚拟化中的重要概念

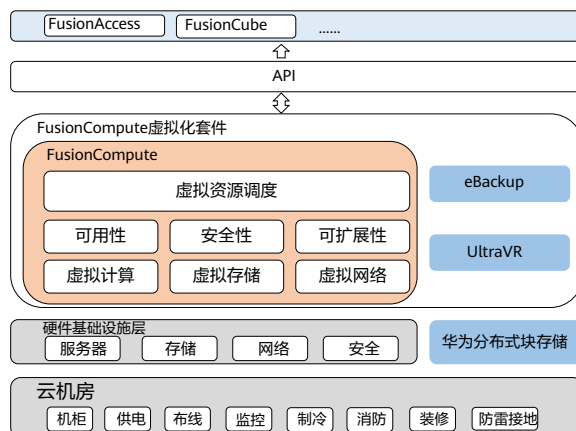


FusionCompute产品定位

- FusionCompute是云操作系统软件，主要负责硬件资源的虚拟化，以及对虚拟资源、业务资源、用户资源的集中管理。它采用虚拟计算、虚拟存储、虚拟网络等技术，完成计算资源、存储资源、网络资源的虚拟化。
- 通过统一的接口，对虚拟资源进行集中调度和管理，从而降低业务的运行成本，保证系统的安全性和可靠性，协助运营商和企业构筑安全、绿色、节能的云数据中心能力。

FusionCompute在虚拟化套件中位置

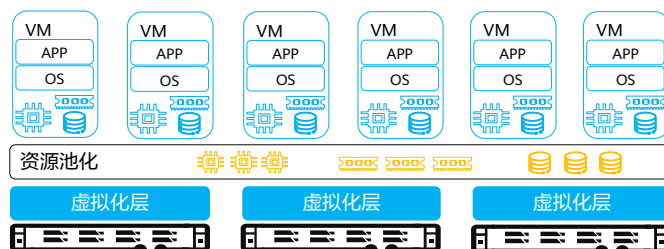
- 云机房
- 硬件基础设施
 - 云计算基础物理设备
 - FusionStorage Block
- FusionCompute虚拟化套件
 - FusionCompute
 - eBackup
 - UltraVR



- 硬件基础设施层：
 - 硬件资源包括服务器、存储、网络、安全等全面的云计算基础物理设备，支持用户从中小规模到大规模的新建或扩容，可运行从入门级到企业级的各种企业应用。设备类型丰富，可为客户提供灵活的部署选择。
- 华为分布式块存储（FusionStorage Block）是一种存储与计算高度融合的分布式存储软件，在通用服务器上部署该软件后，可以把所有服务器的本地硬盘组织成一个虚拟存储资源池，提供块存储功能。
- FusionCompute虚拟化套件：通过在服务器上部署虚拟化软件，将硬件资源虚拟化，从而使一台物理服务器可以承担多台服务器的工作。通过整合现有的工作负载并利用剩余的服务器以部署新的应用程序和解决方案，可以实现较高的整合率。
 - FusionCompute是云操作系统软件，主要负责硬件资源的虚拟化，以及对虚拟资源、业务资源、用户资源的集中管理。它采用虚拟计算、虚拟存储、虚拟网络等技术，完成计算资源、存储资源、网络资源的虚拟化。同时通过统一的接口，对这些虚拟资源进行集中调度和管理，从而降低业务的运行成本，保证系统的安全性和可靠性，协助运营商和企业构筑安全、绿色、节能的云数据中心能力。
 - eBackup是虚拟化备份软件，配合FusionCompute快照功能和CBT（Changed Block Tracking）备份功能实现FusionCompute的虚拟机数据备份方案。
 - UltraVR是容灾业务管理软件，利用底层SAN存储系统提供的异步远程复制特性，提供虚拟机关键数据的数据保护和容灾恢复。

FusionCompute产品功能 - 虚拟化计算

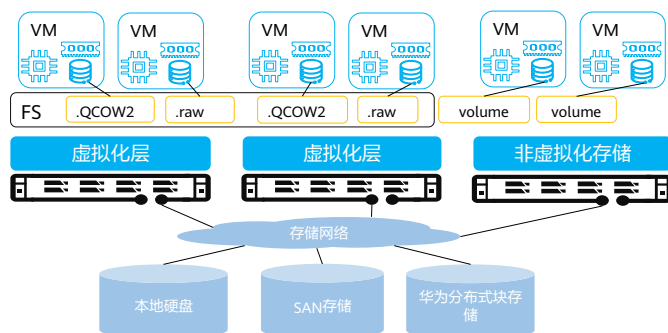
- FusionCompute支持将服务器物理资源抽象成逻辑资源，让一台服务器变成几台甚至上百台相互隔离的虚拟服务器，不再受限于物理上的界限，让CPU、内存、磁盘、I/O等硬件变成可以动态管理的“资源池”，从而提高资源的利用率，简化系统管理。
- 对于最终用户，虚拟机比物理机的优势在于它可以很快速地发放，很方便地调整配置和组网。对于维护人员来讲，虚拟机复用了硬件，这样硬件更少，加上云平台的自动维护能力，维护成本显著降低。对于系统管理员，可以很直观地知道资源使用的总量及变化趋势，以便决策是否扩容。



- 将服务器物理资源抽象成逻辑资源，让一台服务器变成几台甚至上百台相互隔离的虚拟服务器，不再受限于物理上的界限，而是让CPU、内存、磁盘、I/O等硬件变成可以动态管理的“资源池”，从而提高资源的利用率，简化系统管理。同时硬件辅助虚拟化技术提升虚拟化效率，增加虚拟机的安全性。

FusionCompute产品功能 - 虚拟化存储

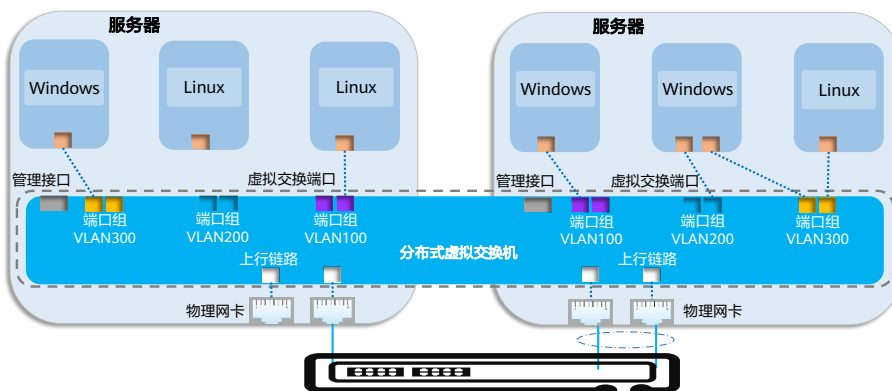
- FusionCompute存储虚拟化技术可以更好地管理虚拟基础架构的存储资源，使系统大幅提升存储资源利用率和灵活性，提高应用的正常运行时间。FusionCompute支持将SAN设备、计算节点本地硬盘以及华为分布式块存储提供的虚拟存储空间统一管理，以虚拟卷的形式分配给虚拟机使用。



- 存储设备的能力、接口协议等差异性很大，存储虚拟化技术可以将不同存储设备进行格式化，将各种存储资源转化为统一管理的数据存储资源，可以用来存储虚拟机磁盘、虚拟机配置信息、快照等信息。用户对存储的管理更加同质化。
- 虚拟机磁盘、快照等内存均以文件的形式存放在数据存储上，所有业务操作均可以转化成对文件的操作，操作更加直观、便捷。
- 基于存储虚拟化平台提供的众多存储业务，可以提高存储利用率，更好的可靠性、可维护性、可以带来更好的业务体验和用户价值。
- 华为提供基于主机的存储虚拟化功能，用户不需要再关注存储设备的类型和能力。存储虚拟化可以将存储设备进行抽象，以逻辑资源的方式呈现，统一提供全面的存储服务。可以在不同的存储形态，设备类型之间提供统一的功能。
- 对于最终用户，就像x86服务器使用本地硬盘一样的方式使用，可以格式化，安装文件系统，安装操作系统，读写。同时，虚拟化的存储还具备快照能力，可以调整大小，这是物理硬盘不能实现的。
- 对于管理员来说，虚拟存储卷并不一对一映射到某块具体的磁盘，而是收敛到几台SAN设备。由于SAN设备已经有了可靠性保障，所以更换硬盘的工作量大规模下降。同时，虚拟存储具备瘦分配，灵活调整，QoS可限制，可迁移等等比物理盘强的特性，在整体成本方面优势很明显。

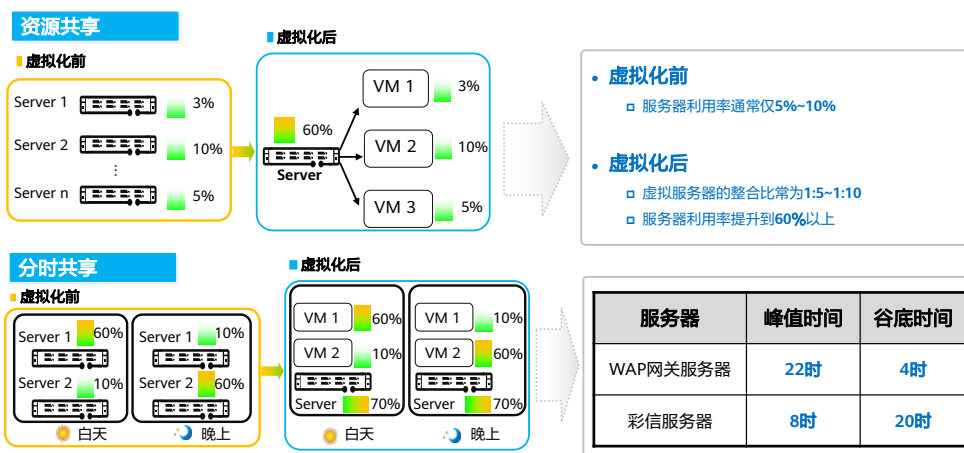
FusionCompute产品功能 - 虚拟化网络

- FusionCompute支持分布式虚拟交换，可以向虚拟机提供独立的网络平面。像物理交换机一样，不同的网络平面间通过VLAN进行隔离。



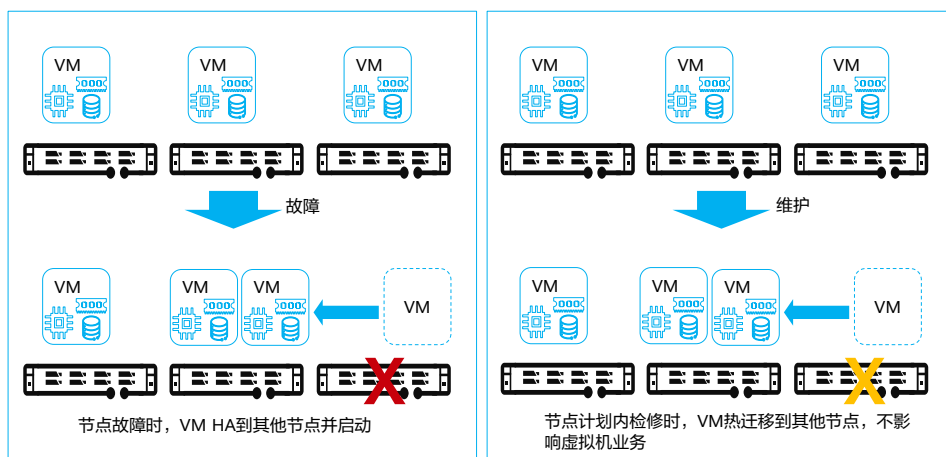
- 分布式交换机的功能类似于普通的物理交换机，每台主机都连接到分布式交换机中。分布式交换机的一端是与虚拟机相连的虚拟端口，另一端是与虚拟机所在主机上的物理以太网适配器相连的上行链路。通过它可以连接主机和虚拟机，实现系统网络互通。另外，分布式交换机在所有关联主机之间作为单个虚拟交换机使用。此功能可使虚拟机在跨主机进行迁移时确保其网络配置保持一致。

FusionCompute带来的价值（1）



- 资源共享：
 - 虚拟化前，需要3台物理机器运行独立应用，每台物理服务器资源利用率都很低；投资回报率低。
 - 虚拟化后，3个独立应用运行在一台服务器上面，需购置的服务器大大减少；且服务器的资源使用率明显提高；投资回报率高。
- 分时共享：
 - 虚拟化前，由于应用运行于不同服务器上，其中一个应用业务压力大时，另一台业务压力小的服务器富裕的计算资源不能转移给这台服务器；从整体来看，服务器资源使用不均衡，资源利用不合理。
 - 虚拟化后，应用共享一个服务器资源；其中一个应用业务压力大，另外一台压力小的情况下，FusionCompute会协助服务器动态地将资源合理分配给各个应用，不仅能满足业务压力的需求，而且保持服务器资源处于高利用率状态。
- 通过资源共享和分时共享解释，回归本页主题，虚拟化能够明显提高资源利用率，降低投资成本并提高投资回报率。

FusionCompute带来的价值（2）



- 后面会提到FusionCompute提供的虚拟化HA功能。
- 当业务虚拟机所在的主机发生故障或需要人工维护时，FusionCompute能够第一时间检测到异常，并把异常机器上运行的虚拟机在集群中其他正常的主机上面重新启动，保障业务连续性。
- FusionCompute通过智能算法还能够维持主机间负载的均衡；从而提供整体的可靠性。
- 总之，通过使用和合理配置FusionCompute，可极大提高系统及运营业务的可靠性。

云操作系统对于服务器的需求

- 云操作系统能够支持计算、网络以及存储的虚拟化，同时在可用性上也支持虚拟机热迁移和故障迁移，在网络安全上通过设置虚拟网卡划分不同的VLAN ID来实现网络管理，而为了帮助云操作系统实现这些功能，服务器硬件也需要满足一定的要求。
- 以部署FusionCompute物理服务器为例，重点要求如下：
 - CPU：CPU需要支持硬件虚拟化技术，同一集群内服务器的CPU为同一型号。
 - 内存：至少大于8G，推荐大于等于48G。
 - 硬盘：在资源管理节点VRM上的系统盘大于等于270G，计算节点大于等于150G。
 - 网口：NIC网口数目大于等于1，建议网卡数目6个，网卡速率千兆及以上。

- 其他主机配置要求如下：
 - CPU：
 - x86：Intel的64位CPU，Hygon的64位CPU；Arm：鲲鹏920处理器，Phytium的64位CPU。
 - x86架构中，CPU需支持硬件虚拟化技术，如Intel的VT-x，并已在BIOS中开启CPU虚拟化功能。
 - 内存：
 - 如果主机用于部署管理节点虚拟机，则主机的内存规格需至少满足管理节点虚拟机内存规格与管理节点虚拟机所在主机的管理域内存规格之和。
 - 华为服务器内存需要按照推荐配置安装，否则可能会导致系统性能达不到最佳效果。
 - 硬盘：
 - 对于计算节点主机，系统盘要求两个SAS盘组RAID1。
 - 如果业务虚拟机使用本地存储，业务虚拟机所使用本地存储需要单独规划，建议本地磁盘组RAID1。
 - 网口：
 - 采用GE组网或10GE组网应取决于客户对网络流量的预估。建议网络负载不超过带宽的60%。
- 除此之外还有其他的较为复杂的配置要求，本页就不赘述了。

目录

1. 计算系统架构
2. 计算系统硬件
3. 计算系统软件
- 4. 计算系统解决方案**

HPC解决方案



全栈自主创新

从芯片、网络、存储、基础软件到集群软件，全栈自研，面向高性能计算应用性能全栈调优

高性能

- 高算力，vs国内x86友商领先20%
- 高能效，vs国内x86友商领先40%，HPL效率>80%
- 典型行业应用优化，vs x86友商性能领先15%+

高易用

- 基础软件一键集成部署，提升部署效率
- 生态完善，已覆盖80%核心算力需求，2500+主流软件迁移（含信BioConda社区软件）
- 典型应用构建容器镜像，免编译免调优

鲲鹏HPC特点

- HPC生态以开源为主，商业软件主要在制造、EDA、油气领域
- 通过众智计划、鲲鹏展翅基金、科研创新计划支持400+软件迁移，覆盖80%以上HPC算力需求
- 气象领域全覆盖，科研行业算力需求满足90%以上

生命科学

<生物基因>

GATK BWA
Canu bowtie2
Blast fasta
Repertmasker phylip

<冷冻电镜>

Emman relion

<药物研发>

gromacs AutoDock

<分子动力学>

gromacs amber
NAMD LAMMPS

制造仿真

<碰撞+NVH>

Radioss Ls-Dyna
Abaqus pam-crash

<电磁仿真>

HFSS Maxwell
AnsysEM Matlab

<结构仿真>

Abaqus COMSOL
ANSYS Mechanical

<流体仿真>

Fluent OpenFoam
Star-CCM+ CFX

气象环境

<海洋模拟>

FVCOM NEMO
ROMS POM
WAVEWATCH III

<气象预测>

GRAPES WRF

<气候预测>

CESM BCC

<空气质量>

CMAQ CAMx
SMOKES WRF-CHEM

科研超算

<计算化学>

GROMACS
LAMMPS VASP

<高能物理>

ROOT GEANT4

<天文物理>

GADGET2

<超算平台>

Gaussian GROMACS
NAMD VSAP
Amber LAMMPS

生态特点: 开源为主
生态进展: 完成1000+软件迁移，覆盖80%算力需求

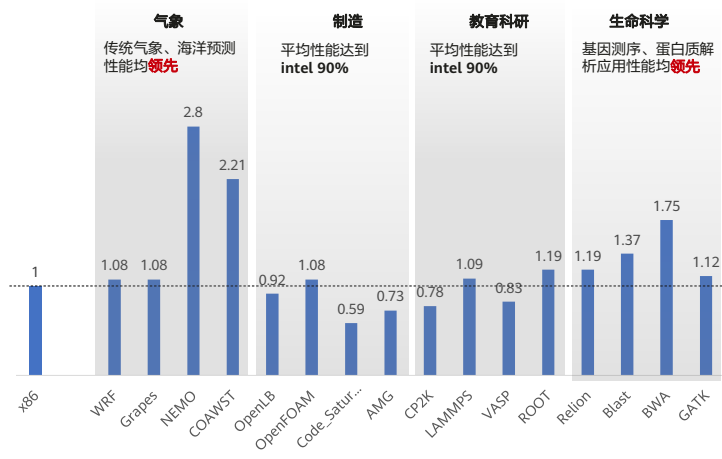
生态特点: 商业软件为主
生态进展: 完成3款原生使能，20多软件二进制转码使能

生态特点: 开源为主，生态聚焦
生态进展: 完成30+主流软件迁移，生态覆盖99%

生态特点: 开源为主
生态进展: 完成400+软件迁移，覆盖90%算力需求

48 Huawei Confidential

高性能鲲鹏HPC



典型应用性能归一化对比

- Intel 8458
- 鲲鹏920高性能

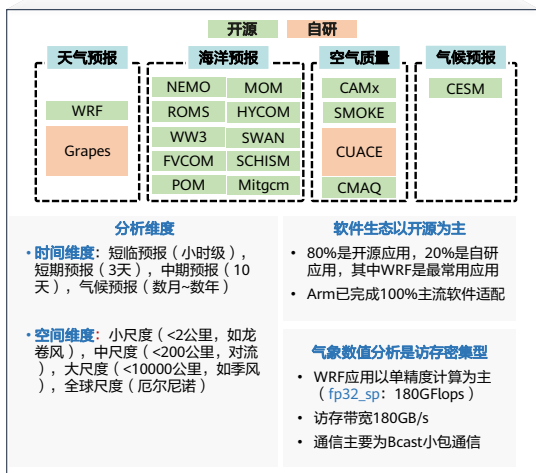
- 测试采用相同节点数
- 来源华为实验室测试数据，结果在不同环境中可能有偏差，具体测试配置如下

服务器	2U2P	2U2P
CPU	鲲鹏920高性能	Intel 8458P
OS	Kylin V10	Centos7.9
Memory	16*32GB 1R DDR4 4800	
Storage	DPC(3.4P)	
MPI	hmpi 1.2.0	oneAPI2021/openmpi 4.1.2
Compiler	bisheng 2.4.0	oneAPI2021/gcc 7.3.0

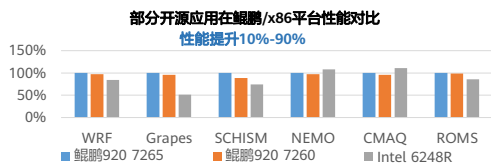


鲲鹏HPC应用场景：气象分析

气象软件生态100%打穿，精度高、时效高



鲲鹏在气象领域成熟规模应用



数学库优化，编译调优，热点代码优化，并行与通信调优，应用代码优化

华北某气象局

- 准确性达到**99.98%+**，同节点性能超过气象局现网X86平台

华东某省气象局

- WRF预报结果在鲲鹏与X86平台达到**逐小时完全一致**，相似度为1，差值全部为0

华南某省气象局

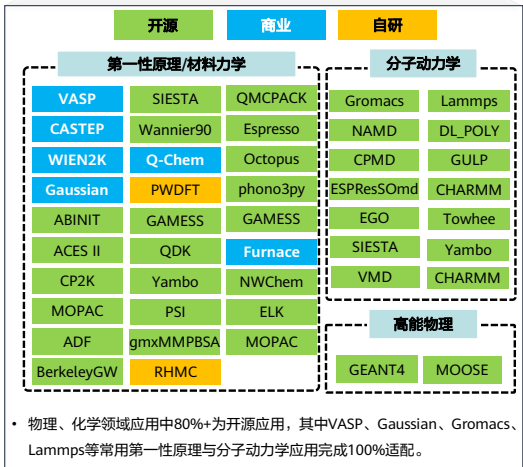
- 9km区域7天台风气象预报仿真计算，相比原来平台**用时减少11%**

华东某省气象局

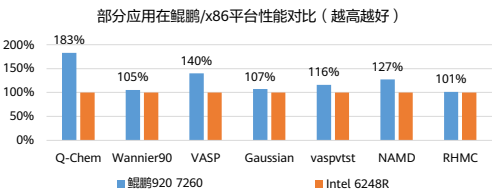
- WRF预报结果在鲲鹏与X86平台到**完全一致**
- 三套模式全部**1H内运行完成**

鲲鹏HPC应用场景：教育科研

基本覆盖常用物理化学计算软件生态



软硬协同优化，综合性能领先x86



关键优化

毕昇C++编译器

- 向量化增强
- 访存优化
- 调度优化

KNL数学库

- 向量指令优化
- 算法处理流程优化
- 并行计算优化

HMPI通讯库

- 拓扑感知优化
- 集合通讯优化

上海XX大学

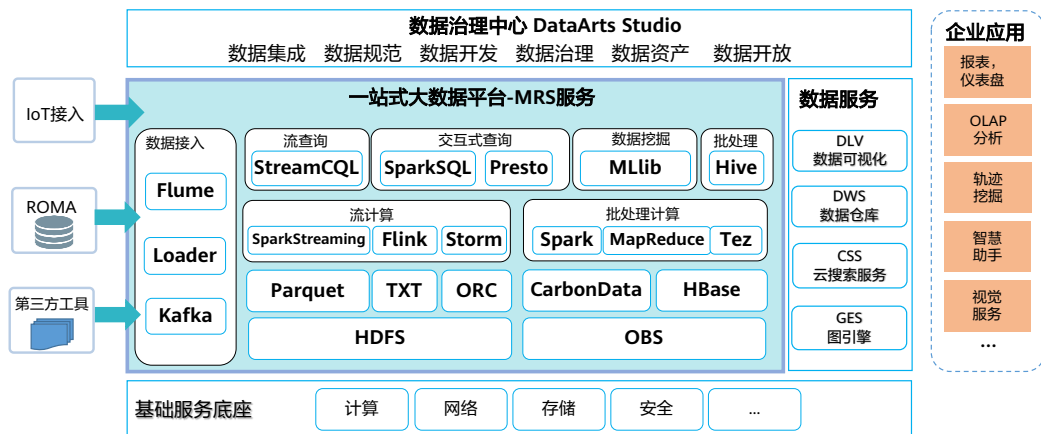
上线基于鲲鹏ARM的校级超算系统，共12800核。服务校内74个课题组，至今计算作业数量超100万个。支撑该校老师的课题组开发的大规模并行分子动力学软件LAMMPS-RBE。

甘肃XX大学

该大学超算中心投入使用鲲鹏HPC集群，该集群具1960计算核心，支撑分子动力学软件GROMACS，LAMMPS，AmberTools、NAMD、CP2k，Quantum Espresso等10余款物理/化学应用。

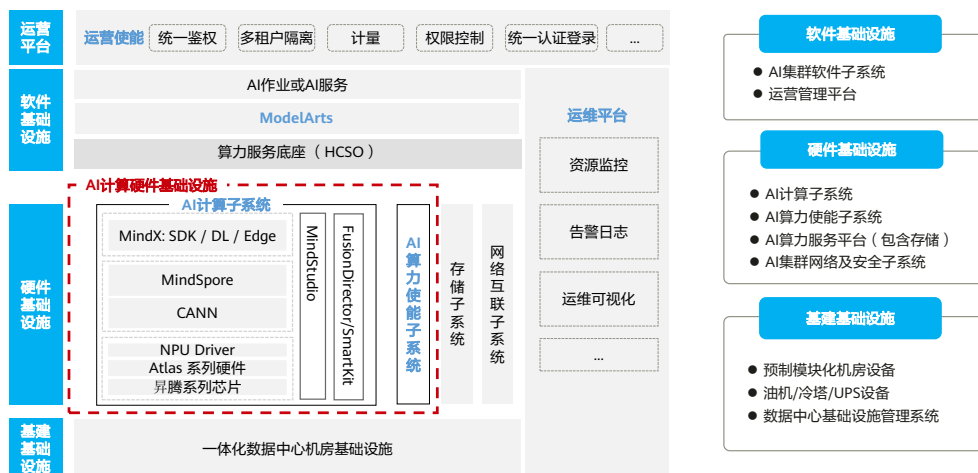
华为云大数据解决方案

- 数据开发，测试，应用一站式服务



- 100%兼容开源生态+三方组件插件化管理，企业一站式平台，支持存算分离+鲲鹏优化，性能更优。
- 围绕着企业数据湖，提供一站式数据资产管理、开发、探索和共享能力，提供统一数据集成、数据治理、开发平台，无缝连接华为云，MRS、DWS、DLI等数据底座的开发。
- 对于数据源复杂、数据类型多样的政务数据，无需切换多个工具实现数据集成工作。

人工智能计算中心解决方案



- 人工智能计算中心解决方案：自主可控、端到端全栈、生态开放。

思考题

1. （判断题）服务器的发展历程为由小型机阶段发展到通用服务器阶段再发展到大型机阶段。
 - A. 正确
 - B. 错误
2. （单选题）以下哪一种数据库模型允许一个以上的节点可以有多于一个的双亲？
 - A. 层次数据模型
 - B. 网状数据模型
 - C. 关系数据模型
 - D. 事务数据模型

• 答案:

- 1. B
- 2. B

思考题

3. （多选题）以下关于计算系统解决方案的描述中，正确的是哪些项？
- A. HPC解决方案硬件平台的存储子系统可以使用OceanStore Pacific产品来搭建
 - B. 在教育和科研领域的应用目前大多是闭源的，鲲鹏HPC生态正在加速这些适配这些软件
 - C. 用户可利用大数据MapReduce服务（MRS）轻松运行HBase、Kafka等大数据组件
 - D. ModelArts是面向开发者的一站式AI开发平台，帮助用户快速创建和部署模型

- 答案：
 - 3. ACD

本章总结

- 通过本章的学习，我们从计算架构全景图出发，首先学习了计算系统硬件，了解了服务器的发展历程，不同的分类方式，以及不同类别服务器各自的特点。
- 然后通过学习计算系统软件，我们了解了以Linux为代表的一些服务器操作系统，对于中间件和数据库，云计算的虚拟化技术和相关产品进行了初步的学习。
- 最后本章着重介绍了华为HPC解决方案，大数据解决方案和AI解决方案，进而了解了鲲鹏和昇腾生态它们的应用场景和优势。

学习推荐

- 鲲鹏社区：
 - <https://www.hikunpeng.com/>
- 昇腾社区：
 - <https://www.hiascend.com/zh/>

缩略语

- CRM: Computing Resource Manager, 计算资源管理器, 主要用于管理集群内的计算节点所映射的虚拟计算资源。
- ERP: Enterprise Resource Planning, 企业资源计划, 一个建立在信息技术基础上, 为企业员工及决策层提供决策手段的管理平台。
- OA: Office Automation, 办公自动化, 是将现代化办公和计算机网络功能结合起来的一种新型的办公方式。
- KVM: Kernel-based Virtual Machine, 内核虚拟机, KVM是一种内建于Linux中的开源虚拟化技术。
- CISC: Complex Instruction Set Computing, 复杂指令集计算机, 一种CPU指令集。
- RISC: Reduced Instruction Set Computing, 精简指令集计算机, 一种CPU指令集。
- DNS: Domain name server, 域名服务器, 在因特网内执行域名服务的功能单元。
- HPC: High-Performance Computing, 高性能计算, 专用于处理大型计算问题的计算机集群系统。
- CPU: Central Processing Unit, 中央处理器, 是计算机的通用计算单元。
- GPU: Graphic Processing Unit, 图形处理器, 执行复杂的数学和几何图形计算的计算单元。

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。
Bring digital to every person, home, and
organization for a fully connected,
intelligent world.

Copyright©2024 Huawei Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.



服务器架构概述



前言

- 通过前面章节的学习，我们了解了什么是服务器，以及服务器类别的划分。那么，服务器是由哪些组件组成的，它们的作用又是什么呢？
- 本章我们将一起来学习服务器的硬件组成、软件、服务器相关的协议，以及管理和控制计算机硬件与软件资源的操作系统。

目标

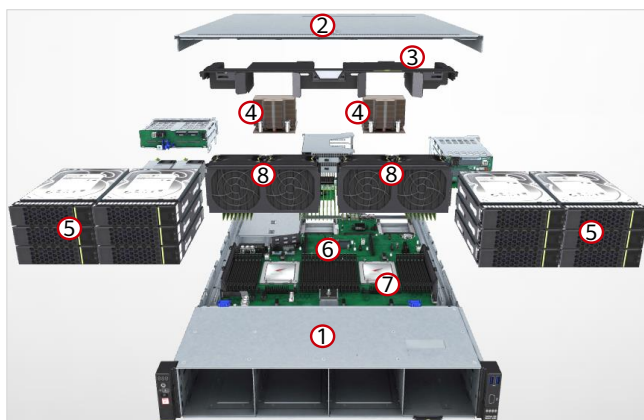
- 学完本课程后，您将能够：
 - 了解服务器的硬件组成，以及它们的功能作用
 - 掌握服务器的常用软件，以及相关的协议
 - 了解操作系统的发展史，以及常用的服务器操作系统

目录

1. 服务器硬件组成
2. 服务器软件介绍
3. 操作系统介绍

服务器关键硬件组成

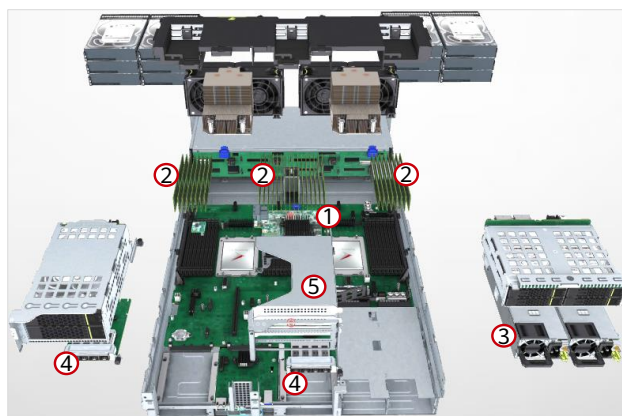
- 以TaiShan200（型号2280）均衡型服务器为例-正面爆炸视图：



- ① 机箱
- ② 机箱盖
- ③ 导风罩
- ④ 散热器
- ⑤ 硬盘
- ⑥ 主板
- ⑦ CPU
- ⑧ 风扇

服务器关键硬件组成

- 以TaiShan200（型号2280）均衡型服务器为例-背面爆炸视图：



- ① RAID卡
- ② 内存
- ③ 电源
- ④ 灵活IO
- ⑤ Riser模组

CPU简介

- CPU（Central Processing Unit，中央处理器），是一块超大规模的集成电路，是一台计算机的运算核心和控制核心。
- 其功能主要是读取、译码并执行计算机指令以及处理计算机软件中的数据。
- CPU、内部存储器、输入/输出设备合称为计算机三大核心部件，由运算器、控制器、寄存器、高速缓存及实现它们之间联系的数据、控制及状态的总线构成。



- CPU：是服务器上的核心处理单元，而服务器是网络中的重要设备，要处理大量的访问需求。因此对服务器具有大数据量的快速吞吐、超强的稳定性、长时间运行等严格要求。所以说CPU是计算机的“大脑”，是衡量服务器性能的首要指标。
- 控制器：是指按照预定顺序改变主电路或控制电路的接线和改变电路中电阻值来控制电动机的启动、调速、制动和反向的主令装置。它是发布命令的“决策机构”，即完成协调和指挥整个计算机系统的操作。
- 运算器：计算机中执行各种算术和逻辑运算操作的部件。运算器的基本操作包括加、减、乘、除四则运算，与、或、非、异或等逻辑操作，以及移位、比较和传送等操作，亦称算术逻辑部件。
- 寄存器：寄存器的主要作用是用来暂时存放参与运算的数据和运算结果，具有接收数据、存放数据和输出数据的功能。

CPU分类

- 一个CPU的体系结构，简而言之，就是包括一组指令集和一些寄存器。常见的CPU体系结构有ARM，MIPS，PowerPC，x86，TI等。
 - x86 CPU架构：
 - CISC（Complex Instruction Set Computer，复杂指令集计算机），其中每个指令可以执行一些较低阶的硬件操作，指令数目多而且复杂，每条指令的长度并不相同。由于指令执行较为复杂所以每条指令花费的时间较长。
 - ARM CPU架构：
 - RISC（Reduced Instruction Set Computer，精简指令集计算机），旨在执行较少类型计算机指令，以便能够以更高的速度执行操作，使计算机的结构更加简单、合理地提高运行速度。

- 鲲鹏CPU架构相对于x86CPU架构具有更加均衡的性能功耗比。早期的CPU全部是CISC架构，它的设计目的是要用最少的机器语言指令来完成所需的计算任务。这种架构会增加CPU结构的复杂性和对CPU工艺的要求，但对于编译器的开发十分有利。RISC架构要求软件来指定各个操作步骤。这种架构可以降低CPU的复杂性以及允许在同样的工艺水平下生产出功能更强大的CPU，但对于编译器的设计有更高的要求。
- 采用RISC架构的ARM微处理器一般具有如下特点：
 - 体积小、低功耗、低成本、高性能；
 - 支持Thumb（16位）/ARM（32位）双指令集，能很好的兼容8位/16位器件；
 - 大量使用寄存器，指令执行速度更快；
 - 大多数数据操作都在寄存器中完成；
 - 寻址方式灵活简单，执行效率高；
 - 指令长度固定。

CPU架构

- **控制单元：**

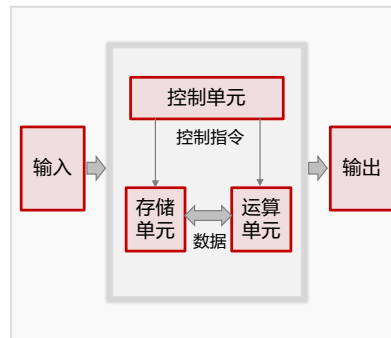
- 控制单元是整个CPU的指挥控制中心，由PC（Program Counter，程序计数器）、IR（Instruction Register，指令寄存器）、ID（Instruction Decoder，指令译码器）和OC（Operation Controller，操作控制器）等组件组成。它依次从存储器中取出各条指令，放在指令寄存器IR中，通过指令译码（分析）确定应该进行什么操作，然后通过操作控制器OC，按确定的时序，向相应的部件发出微操作控制信号。

- **运算单元：**

- 运算单元是运算器的核心，可以执行算术运算和逻辑运算。运算器接受控制单元的命令后执行动作，是执行部件。

- **存储单元：**

- 包括CPU片内缓存和寄存器组，是CPU中暂时存放数据的地方，里面保存着那些等待处理的数据，或已经处理过的数据，CPU访问寄存器所用的时间要比访问内存的时间短。采用寄存器，可以减少CPU访问内存的次数，从而提高了CPU的工作速度。



- CPU的运行原理就是：控制单元在时序脉冲的作用下，将指令计数器里所指向的指令地址（这个地址是在内存里的）送到地址总线上，然后CPU将这个地址里的指令读到指令寄存器进行译码。对于执行指令过程中所需要用到的数据，会将数据地址也送到地址总线，然后CPU把数据读到CPU的内部存储单元（就是内部寄存器）暂存起来，最后命令运算单元对数据进行处理加工。
- 存储器包括：高速缓冲存储器（Cache）和内部存储器（Memory）。

CPU基本功能



CPU关键参数



- 主频即CPU的时钟频率，反应了CPU的运算速度，但并不是一个简单的线性关系。CPU的运算速度还要看CPU的流水线、总线等各方面的性能指标。提高主频对于提高CPU运算速度至关重要。
- 例如，假设某个CPU在一个时钟周期内执行一条运算指令，那么当CPU运行在100 MHz主频时，将它运行在50 MHz主频时速度快一倍。工作在100 MHz主频的CPU执行一条运算指令所需时间仅为10 ns，比工作在50 MHz主频时的20 ns缩短了一半，自然运算速度也就快了一倍。只不过计算机的整体运行速度不仅取决于CPU运算速度，还与其它各分系统的运行情况有关。

- CPU高速缓存（CPU Cache）的出现主要是为了解决CPU运算速度与内存读写速度不匹配的矛盾，提升系统的性能。其容量远小于内存，但速度却可以接近处理器的频率。缓存大小是CPU的重要指标之一，CPU内缓存的运行频率极高，一般是和处理器同频运作，工作效率远远大于系统内存和硬盘。但是从CPU芯片面积和成本的因素来考虑，缓存都很小。

- CPU的核又称为core，增加core可以增加线程，一般情况下它们是1:1对应关系，也就是说16核CPU一般拥有16个线程。但Intel引入超线程技术后，使核心数与线程形成1:2的关系，大幅提升了其多任务、多线程性能。

- CPU核数：单块 CPU 上面能处理数据的芯片组的数量，如双核、四核等。双核，就是将两个计算内核集成在一个硅晶元上。从而提高计算能力。多核就是将多个计算内核集成在一个硅晶元上。

多路CPU

- 多路CPU：
 - “路”是指服务器上物理CPU的数量，也就是服务器主板上CPU插槽的数量。从服务器的主板实际插入CPU的个数，可以看出CPU的路数。
 - 单路：指服务器支持1个CPU；双路：指服务器支持2个CPU；四路：指服务器支持4个CPU；多路：指服务器支持多个CPU。
- 例如：
 - TaiShan 200 服务器（型号2280）支持以下处理器型号：

处理器型号	服务器支持路数	处理器规格
鲲鹏920 7260 处理器	2路	64核 2.6 GHz
鲲鹏920 5250 处理器	2路	48核 2.6 GHz
鲲鹏920 5220 处理器	2路	32核 2.6 GHz
鲲鹏920 3210 处理器	2路	24核 2.6 GHz



- CPU 最初发展的时候是一个 CPU 一个处理核心，CPU 的性能主要靠提高核心工作频率来提高，但是仅仅提高单核芯片的速度会产生过多热量且无法带来相应的性能改善。为了提升处理器的能效，于是发展出来了双核心 CPU（Dual-core processor）和多核心的 CPU（Multi-core processor），在物理上是把 2 个或者更多的独立处理器芯片封装在一个单一的集成电路中。

内存

- 内存定义

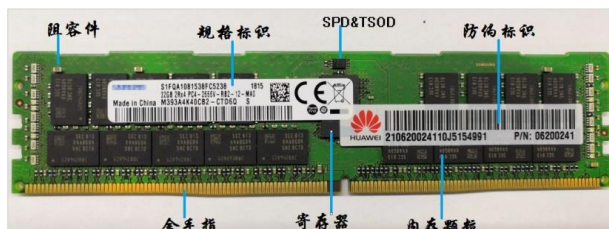
- 内存是服务器中重要的部件之一，它是外存与CPU进行沟通的桥梁。计算机中所有程序的运行都是在内存中进行的，内存的性能对服务器的影响非常大。
- 内存（Memory）也被称为内存储器和主存储器，其作用是用于暂时存放CPU中的运算数据，以及与硬盘等外部存储器交换的数据。
- 只要服务器在运行中，操作系统就会把需要运算的数据从内存调到CPU中进行运算，当运算完成后CPU再将结果传送出来，内存的稳定读写也决定了服务器的稳定运行。



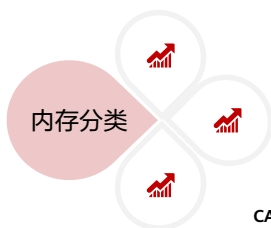
- 在计算机的组成结构中，有一个很重要的部分就是存储器。存储器是用来存储程序和数据的部件，对于计算机来说，有了存储器，才有记忆功能，才能保证正常工作。
- 内存是电脑中的主要部件，它是相对于外存而言的。我们平常使用的程序，如Windows操作系统、打字软件、游戏软件等，一般都是安装在硬盘等外存上的，必须把它们调入内存中运行，才能真正使用其功能。我们平时输入一段文字，或玩一个游戏，其实都是在内存中进行的。就好比在一个书房里，存放书籍的书架和书柜相当于电脑的外存，而我们工作的办公桌就是内存。通常我们把要永久保存的、大量的数据存储在外存上，而把一些临时的或少量的数据和程序放在内存上。
- 内存是计算机中重要的部件之一，它是与CPU进行沟通的桥梁。计算机中所有程序的运行都是在内存中进行的，因此内存的性能对计算机的影响非常大。内存由内存芯片、电路板、金手指等部分组成。
- 服务器内存插槽及配置原则：
 - 同一台服务器必须使用相同型号的DIMM。
 - CPU对应的内存槽位上必须至少配置一根内存条。
 - 当服务器配置完全平衡的内存条时，可实现最佳的内存性能。不平衡配置会降低内存性能，因此不推荐使用。

内存条基本结构

- **内存条基本结构**：阻容件，规格标识，SPD&TSOD，防伪标识，金手指，寄存器，内存颗粒。
 - SPD（Serial Presence Detect，串行存在检测），是一颗8针的EEPROM（Electrically Erasable Programmable ROM，电可擦写可编程只读存储器），容量为256字节，里面主要保存了该内存的相关资料，如容量、芯片厂商、内存模组厂商、工作速度等。SPD的内容一般由内存模组制造商写入。支持SPD的主板在启动时自动检测 SPD中的资料，并以此设定内存的工作参数。
 - 寄存器（Register）：时序校准功能。



内存分类



RAM (Random Access Memory, 随机存储器)

RAM既可以从存储器中读取数据，也可以向存储器中写入数据。当设备断电时，存于其中的数据就会丢失。内存条 (SIMM) 就是将RAM集成块集中在一起的一小块电路板，它插在服务器中的内存插槽上，以减少RAM集成块占用的空间。RAM分为两种：DRAM (Dynamic RAM, 动态随机存储器) 和SRAM (Static RAM, 静态随机存储器)。

ROM (只读存储器, Read Only Memory)

ROM是一种在正常工作时其存储的数据固定不变的存储器，这些数据只能读取，不能写入。即使设备断电，这些数据也不会丢失。ROM一般用于存放服务器的基本程序和数据，如BIOS ROM，其物理外形一般是双列直插式 (DIP) 的集成块。现在比较流行的只读存储器是闪存 (Flash Memory)，它属于EEPROM的升级，可以通过电学原理反复擦写。现在大部分BIOS程序就存储在Flash ROM芯片中。U盘和固态硬盘 (SSD) 也是利用闪存原理做成的。

CACHE (高速缓冲存储器)

Cache也是我们经常遇到的概念，也就是平常看到的一级缓存 (L1 Cache)、二级缓存 (L2 Cache)、三级缓存 (L3 Cache) 这些数据，它位于CPU与内存之间，是一个读写速度比内存更快的存储器。当CPU向内存中写入或读出数据时，这个数据也被存储进高速缓冲存储器中。当CPU再次需要这些数据时，CPU就从高速缓冲存储器读取数据，而不是访问较慢的内存，当然，如需要的数据在Cache中没有，CPU会再去读取内存中的数据。

- DRAM (Dynamic RAM, 动态随机存储器) 的存储单元是由电容和相关元件做成的，电容内存储电荷的多少代表信号0和1。电容存在漏电现象，电荷不足会导致存储单元数据出错，所以DRAM需要周期性刷新，以保持电荷状态。DRAM结构较简单且集成度高，通常用于制造内存条中的存储芯片。
- SRAM (Static RAM, 静态随机存储器) 的存储单元是由晶体管和相关元件做成的锁存器，每个存储单元具有锁存“0”和“1”信号的功能。它的速度快且不需要刷新操作，但集成度差和功耗较大，通常用于制造容量小但效率高的CPU缓存。
- ROM，传统意义上，指的确实是只能读取不能写入的存储器，例如在早期的计算机系统中，操作系统或者固件常常被烧录在这种ROM中。既然它被烧录进去，内容就不能再次修改，所以被称为“只读存储器”。然而，随着技术的进步，现在的“ROM”通常指的是可以被程序重写但不会因为掉电而丢失数据的非易失存储器。这种存储器包括了EPROM (可擦写可编程只读存储器)、EEPROM (电可擦写可编程只读存储器) 和闪存 (Flash Memory) 等。虽然这些存储器的内容可以被修改，但它们仍然被称为ROM，主要是由于它们的用途仍然与原始的ROM类似：它们通常被用来存储固件或其他在系统启动时需要读取但不常改变的数据。所以，尽管“只读存储器”这个名字可能给人一种误导，但由于历史原因和它们的用途，我们仍然将这些可以被重写的非易失存储器称为ROM。

内存相关概念



内存频率



内存频率，是指内存的运行速度，是衡量内存性能的关键指标之一，通常以MT/s作为计量单位。常见的内存频率有2666、3000、3200、3600。

DIMM



DIMM (Dual Inline Memory Module, 双列直插内存模块)，是指引脚插槽，也就是物理结构方面的分类。DIMM是在单列直插存储器模块 (single inline memory module, SIMM) 的基础上发展起来的，SIMM提供32位数据通道，而DIMM则提供了64位的数据通道。

SDRAM



SDRAM (Synchronous Dynamic Random-Access Memory, 同步动态随机存取存储器)，是一种利用同步计时器对存储器的输出信号加以控制的动态随机存储器。

DDR5

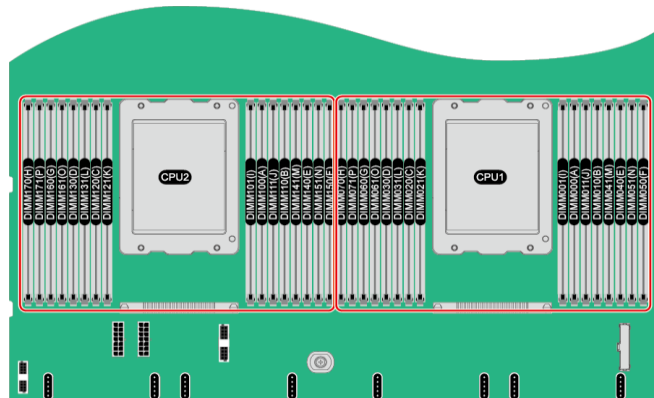


DDR5 SDRAM (Double-Data-Rate Fifth Generation SDRAM, 第五代双倍数据率同步动态随机存储器)，是一种高带宽的内存规格。

- MT/s，即百万次每秒，是内存标准化传输速率的单位，表示每秒的传输次数。

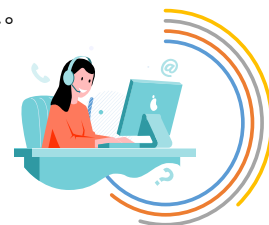
内存槽位编号示例

- 以TaiShan 200服务器（型号2280）为例。当配置2个鲲鹏920 7260或5250处理器时，服务器最大提供32个内存插槽，1个处理器提供8个内存通道，每个通道支持2个DIMM。



内存安装原则

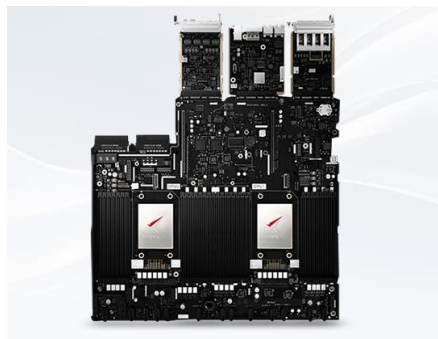
- CPU1对应的内存槽位上至少配置一根内存。
- 同一台服务器不允许混合使用不同规格（容量、位宽、rank、高度等）的内存。即一台服务器配置的多根内存条必须为相同Part No.（即P/N编码）。
- 当服务器配置完全平衡的内存时，可实现最佳的内存性能。不平衡配置会降低内存性能，因此不推荐使用。
- 不平衡的内存配置是指安装的内存不是均匀分布在内存通道或处理器上。
 - **通道不平衡**：如果单个CPU配置3、5、7、9、10、11、12、13、14、15根内存，则通道之间的内存配置不平衡。
 - **处理器不平衡**：如果在每个处理器上安装了不同数量的内存，则处理器之间的内存配置不平衡。



- 内存配置助手链接<https://info.support.huawei.com/computing/smca/?language=zh>

主板

- 主板是服务器的主电路板。
 - 主板的主要功能是提供一系列接合点，形成一个能集成处理器、内存、存储设备（硬盘、固态驱动器、闪存等）、显卡、声卡、网卡和各种外部设备的连接平台。
 - 主板通常直接固定在机箱内壁，表面设计有可以直接插入扩展卡的插槽或可以连接其它内置设备（比如光驱）的线路连接。
 - 主板上最重要的构成组件是**芯片组**。芯片组为主板提供一个通用平台供不同设备连接，控制不同设备的沟通。它亦包含对不同扩展插槽的支持，例如处理器、PCI Express。芯片组亦为主板提供额外功能，例如内置于芯片组的GPU。



- 以鲲鹏服务器主板（型号：S920X00）为例：搭载2颗鲲鹏920处理器，Hi1710智能管理芯片，最多支持32个DDR4-2933 内存插槽、8个PCIe 4.0插槽，适配2U或4U高度机箱的机架服务器，适用于大多数通用数据处理场景。

机框

- 服务器尺寸

- 在服务器领域，1U、2U、4U是常用的服务器尺寸标准，它们指的是服务器机箱的高度，通常用“U”（Unit）作为单位，1U等于1.75英寸（约为4.45厘米）的高度，因此2U就是3.5英寸（约为8.9厘米），4U就是7英寸（约为17.8厘米）。
- 例如，TaiShan 200服务器（型号2280）是2U 2路机架服务器。



前面板

- 同一服务器可以配置不同类型的硬盘。以TaiShan 200服务器（型号2280）为例。可以配置12x3.5英寸硬盘，25x2.5英寸硬盘，8x2.5 SAS/SATA + 12x2.5 NVMe硬盘，24x2.5 SAS/SATA硬盘，或配置8x2.5英寸硬盘。
- 下图以TaiShan 200服务器（型号2280）为例，配置12x3.5英寸硬盘为例展示前面板组件：



① UID 按钮/指示灯

② 健康状态指示灯

③ 电源按钮/指示灯

④ 故障诊断数码管

⑤ 灵活IO卡在位指示灯

⑥ 硬盘

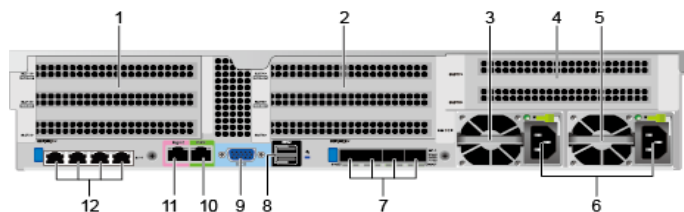
⑦ VGA接口

⑧ USB 3.0 接口

⑨ 标签卡（含SN标签）

后面板

- 服务器的后面板提供了各种模组及接口。下图以TaiShan 200服务器（型号2280）为例：



1	IO模组1	2	IO模组2
3	电源模块1	4	IO模组3
5	电源模块2	6	电源模块接口
7	灵活IO卡2（归属CPU 2）	8	USB 3.0接口
9	VGA接口	10	调试串口
11	管理网口	12	灵活IO卡1（归属CPU 1）

说明：IO模组1、IO模组2和IO模组3都可选配硬盘模组或者Riser模组。

硬盘简介

- 硬盘是计算机最主要的存储设备。
- 硬盘接口是硬盘与主机系统间的连接部件，作用是在硬盘缓存和主机内存之间传输数据。不同的硬盘接口决定着硬盘与计算机之间的连接速度，直接影响着程序运行快慢和系统性能好坏。

硬盘类型	机械硬盘			固态硬盘
	SATA	SAS	NL-SAS	SSD
主流转速 (RPM)	7,200	15,000/10,000	7,200	NA
串行/并行	串行	串行	串行	串行
主流容量 (TB)	1 T/2 T/3 T	0.6 T/0.9 T/1.8 T/2.4 T	2 T/3 T/4 T/8 T/10 T	480 G/960 G/1.92 T/3.84 T
MTBF (h)	1,200,000	1,600,000	1,200,000	2,000,000
备注	由ATA硬盘发展而来，采用串行的方式传输，SATA3.0实现600MB/s最高数据传输率。 SATA硬盘年故障率大约2%。	SAS专为满足高性能企业需求而设计，并且兼容SATA硬盘。能够提供3.0Gb/s到6.0Gb/s的传输率，未来规划到12.0Gb/s。 SAS硬盘年故障率小于2%。	带有SAS接口的“企业级SATA驱动器”，适用于在一个磁盘阵列系统中实现分级存储，简化了磁盘阵列系统的设计。 NL-SAS硬盘年故障率大约2%。	固态硬盘是用固态电子存储芯片阵列而制成的硬盘，由控制单元和存储单元（FLASH芯片、DRAM芯片）组成。固态硬盘的接口规范和定义、功能及使用方法与普通硬盘完全相同，在产品外形和尺寸上也完全与普通硬盘一致。

- MTBF: MeanTimeBetweenFailure，平均无故障工作时间。
- 在价格方面，通常来说SATA和NL-SAS硬盘较为便宜，SAS硬盘较贵一些，SSD最为昂贵。

硬盘的业务类型

硬盘	按业务分类			
HDD	企业级Performance类 这类硬盘不仅性能高，可靠性也很好，但因转速较高，功耗也比同尺寸的其他类型硬盘高。	企业级Capacity类 这类硬盘容量较大，同系列硬盘既有SATA接口，也有SAS接口。	企业级云盘类 主要优势为低功耗、低成本且能实现7x24小时业务应用。	桌面级硬盘类 主要包含PC应用的硬盘，可靠性比企业级低，成本低廉，一般较适合5x8小时业务应用，家用一般选择此类。

硬盘	按业务分类		
SSD	读密集型 (Read intensive) DWPDP (Drive Write Per Day, 每日整盘写入次数) 小于3, 主要适合读业务较频繁的业务。	写密集型 (Write Intensive) DWPDP为10以上, 适合高频度写的业务场景。	读写混合型 (Mixed Use) DWPDP为3~10左右, 适合读写均衡的场景。

- DWPDP全称为Drive Writes Per Day，即每日全盘写入次数，指在预期寿命内可每日完整写入SSD固态硬盘所有容量的次数。在企业级固态硬盘产品里DWPDP这个参数比较常用。

硬盘的接口分类

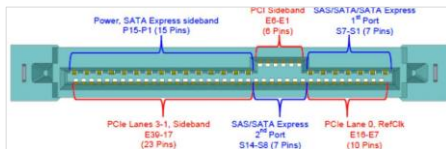
- 当前主流的硬盘使用的接口有SATA、SAS、PCIe（NVMe）等。
- SAS和SATA接口区分：通过红圈2区分，SATA 硬盘红圈2部分是分开的；SAS是连接在一起的，且背面有金手指。



SATA接口



SAS接口



PCIe接口

SAS/SATA 接口最大传输速率关系表

接口类型/协议版本	1.0	2.0	3.0
SAS	3 Gb/s	6 Gb/s	12 Gb/s
SATA	1.5 Gb/s	3 Gb/s	6 Gb/s

说明：PCIe各版本传输速率请查看PCIe总线内容。

什么是NVMe

- NVMe (Non-Volatile Memory Express, 非易失性存储器标准) 是高性能PCIe SSD产品的驱动标准, 主要目的是统一业界各厂商PCIe SSD的驱动程序, 并集成到操作系统(OS)中。对外呈现的接口是PCIe。
- NVMe的优势:
 - 性能高: 通过PCIe直连CPU, 高IOPS、高带宽、低时延。
 - 标准化: OS生态成熟。

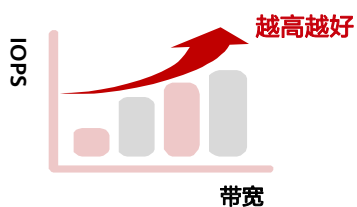


硬盘的尺寸

- HDD按照碟片的尺寸大小，即硬盘碟片直径分为：1-inch、1.8-inch、2.5-inch、3.5-inch、5.25-inch。SSD虽然没有碟片概念，但尺寸大小和HDD是一致的。
- HDD主要尺寸
 - 主流的硬盘中，有3.5寸和2.5寸两种，不同硬盘厂商的外形尺寸都在公差范围内，差异极小。
 - 3.5寸：LFF，长*宽=147 mm*101.85 mm，企业级3.5寸硬盘的厚度一般为26.1 mm。
 - 2.5寸：SFF，长*宽=100.45 mm*69.85 mm，企业级2.5寸硬盘的厚度一般为15 mm。
- SSD主要尺寸
 - 主流的固态硬盘尺寸为2.5寸，不同硬盘厂商的外形尺寸差异微小。
 - 2.5寸：SFF，长*宽=100.45 mm*69.85 mm，SSD的高度有7 mm，9.5 mm和15 mm，各个厂家可能设计不同的厚度，但不会超过15 mm。

硬盘关键指标

- 硬盘容量 (Volume)
- 转速 (Rotational Speed) (机械硬盘指标)
- 平均访问时间 (Average Access Time)
- 数据传输率 (Data Transfer Rate)
- IOPS (Input/Output Operations Per Second)



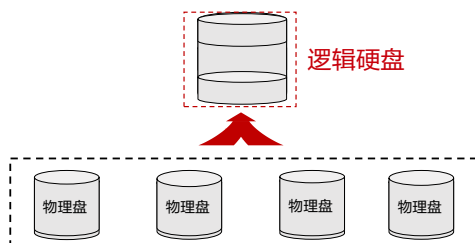
硬盘类型	IOPS (4K随机写)	带宽 (128K顺序读)
SATA	330	200 MB/s
SAS 10K	350	195 MB/s
SAS 15K	450	290 MB/s
SATA SSD	30~60K	540 MB/s
SAS SSD	155K	1000 MB/s
NVMe SSD	300K	3500 MB/s

RAID定义

- RAID定义：

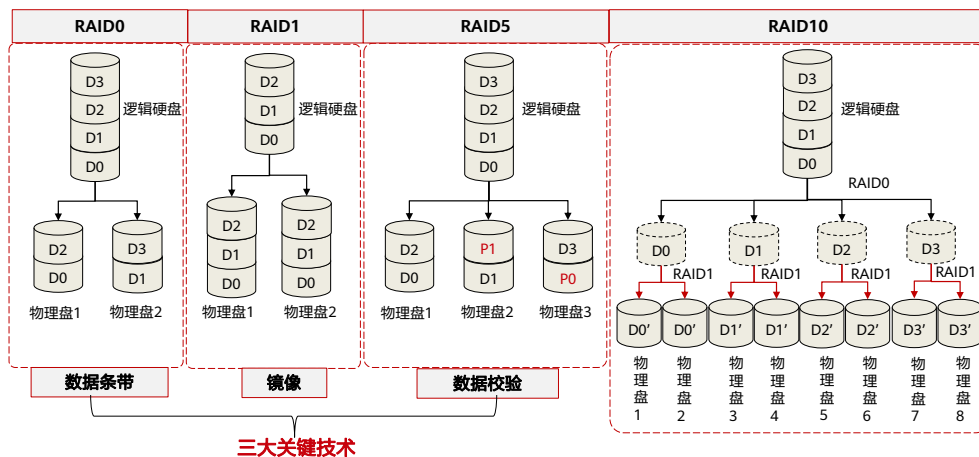
- RAID（Redundant Array of Independent Disks，独立硬盘冗余阵列），旧称廉价磁盘冗余阵列（Redundant Array of Inexpensive Disks），简称磁盘阵列。利用虚拟化存储技术把多个硬盘组合起来，成为一个或多个硬盘阵列组，目的为提升性能或数据冗余，或是两者同时提升。

— 维基百科



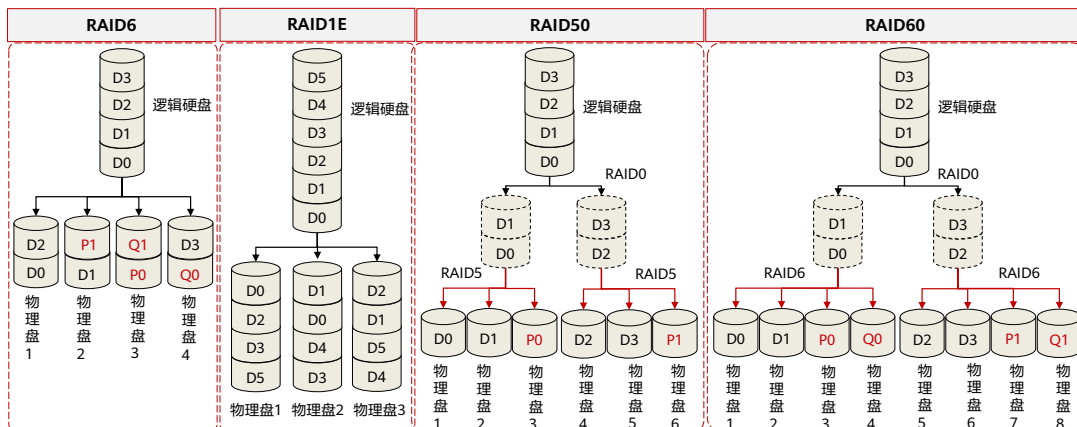
- RAID为廉价磁盘冗余阵列（Redundant Array of Inexpensive Disks），RAID技术将一个个单独的磁盘以不同的组合方式形成一个逻辑硬盘，从而提高了磁盘读取的性能和数据的安全性。不同的组合方式用RAID级别来标识。RAID技术是由美国加州大学伯克利分校D.A. Patterson教授在1988年提出的，作为高性能、高可靠的存储技术，在今天已经得到了广泛的应用。
- RAID技术经过不断的发展，现在已拥有了从 RAID0到RAID5等6种明确标准级别的RAID级别。另外，其他还有6、7、10（RAID1与RAID0的组合）、01（RAID0与RAID1的组合）、30（RAID3与RAID0的组合）、50（RAID0与RAID5的组合）等。不同RAID级别代表着不同的存储性能、数据安全性和存储成本。

RAID级别详解（1）



- 在各个raid级别中，使用最广泛的是RAID0，RAID1，RAID10，RAID5。
- RAID0，将数据分成条带顺序写入一组磁盘中。RAID0不提供冗余功能，但是它却提供了卓越的吞吐性能，因为读写数据是在一组磁盘中的每个磁盘上同时处理的，吞吐性能远远超过单个磁盘的读写。
- RAID1，每次写操作都将分别写两份到数据盘和校验盘上，每对数据盘和校验盘成为镜像磁盘组。也可使用并发的方式来读数据时，提高吞吐性能。如果镜像磁盘组中某个磁盘出错，则数据可以从另外一块磁盘获得，而不会影响系统的性能，然后，使用一块备用磁盘将健康磁盘中的数据复制出来然后这两块磁盘又组成新的镜像组。
- RAID10，即RAID1与RAID0的结合，既做镜像又做条带化，数据先镜像再做条带化。这样数据存储既保证了可靠性，又极大地提高了吞吐性能。
- RAID01也是RAID0与RAID1的结合，但它是对条带化后的数据进行镜像。但与RAID10不同，一个磁盘的丢失等同于整个镜像条带的丢失，所以一旦镜像盘失败，则存储系统成为一个RAID0系统（即只有条带化）。
- RAID5 是将数据校验循环分散到各个磁盘中，它像RAID0一样将数据条带化分散写到一组磁盘中，但同时它生成校验数据做为冗余和容错使用。校验磁盘包含了所有条带的数据的校验信息。RAID5将校验信息轮流地写入条带磁盘组的各个磁盘中，即每个磁盘上既有数据信息又同时有校验信息，RAID5的性能得益于数据的条带化，但是某个磁盘的失败却将引起整个系统的下降，这是因为系统将在承担读写任务的同时，重新构建和计算出失败磁盘上的数据，此时要使用备用磁盘对失败磁盘的数据重建恢复整个系统的健康。

RAID级别详解（2）



- RAID6是指带有两种分布存储的奇偶校验码（既P和Q）的独立硬盘结构。与RAID5相比，RAID6增加了第二个独立校验码（Q）信息块，两个独立的奇偶校验系统使用不同的算法，数据的可靠性非常高，即使两块硬盘同时失效也不会影响数据的使用，主要是用于要求数据绝对安全的场合。
- RAID6的P校验和RAID5的校验是一样的，都是采用的“异或”运算。异或运算如下： $P0 = D0 \text{ xor } D1, Q0 = (D0 * K0) \text{ xor } (D1 * K1)$ （“*”号代表伽罗瓦域乘，K为常量系数，取值1到255之间正整数）
- RAID1E的各个物理盘中，以两个条带组为一个单元，在一个单元中，上一个条带组内是连续排列的数据，下一个条带组则是上一个条带组的错位镜像，即将上一个条带组中每个成员盘内的数据向右移动一块盘的位置写入下一个条带组的各个成员盘中，最后一个成员盘的数据则写入第一个成员盘中。
- RAID1E在一块成员盘离线的情况下不会影响数据的完整性，如果RAID-1E由四块或者四块以上成员盘构成，则能够支持在两块成员盘离线的情况下不会影响数据的完整性，只是有两个前提：一是离线的两块成员盘不能是相邻的两块盘；二是第一块成员盘和最后一块成员盘不能同时离线。
- RAID50是RAID5与RAID0的组合，先作RAID5，再作RAID0，也就是对多组RAID5彼此构成Stripe访问。由于RAID50是以RAID5为基础，而RAID5至少需要3颗硬盘，因此要以多组RAID5构成RAID50，至少需要6颗硬盘。
- RAID60是RAID6与RAID0的组合，先作RAID6，再作RAID0。RAID6至少需具备4颗硬盘，所以RAID60的最小需求是8颗硬盘。

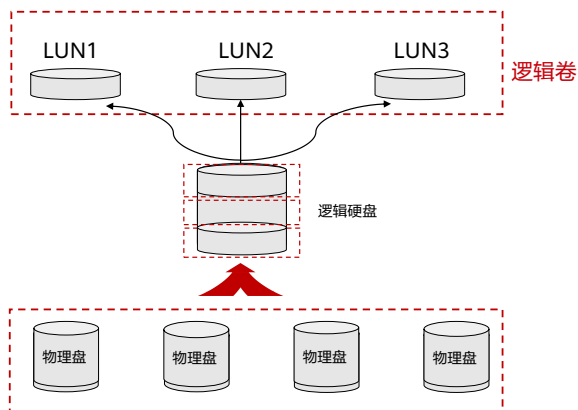
常见RAID级别的比较

RAID级别	可靠性	读性能	写性能	硬盘利用率
RAID0	低	高	高	100%
RAID1	高	高	中	50%
RAID5	较高	高	中	$(N-1)/N$
RAID6	较高	高	中	$(N-2)/N$
RAID10	高	高	中	50%
RAID1E	高	中	中	50%
RAID50	高	高	较高	$(N-M)/N$
RAID60	高	高	较高	$(N-M*2)/N$
注：N为RAID成员盘的个数，M为RAID的子组数。				

• 应用场景：

- RAID0：需要高读/写速度，但安全性低的环境，如图形工作站。
- RAID1：需要随机数据写入和高安全性的保证，如服务器、数据库、存储等。
- RAID1E：需要数据传输和高安全性场景，如视频编辑和大型数据库存储等。
- RAID5/6：需要一定的随机数据传输和较高安全性，如金融和数据库存储等场景。
- RAID10：对随机读写和安全性要求高的场景，如银行和金融等。

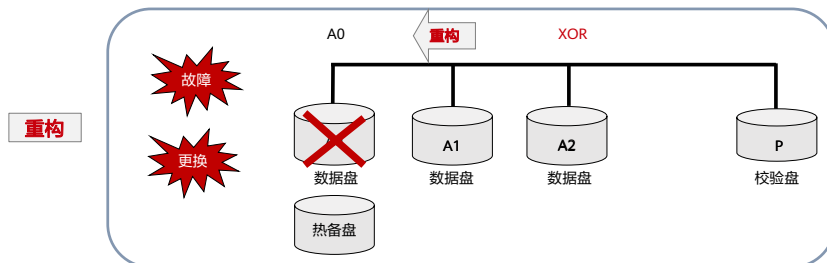
RAID逻辑卷概念



- 在RAID的基础上可以按照指定容量创建一个或多个逻辑卷，通过LUN（Logic Unit Number）来标识。

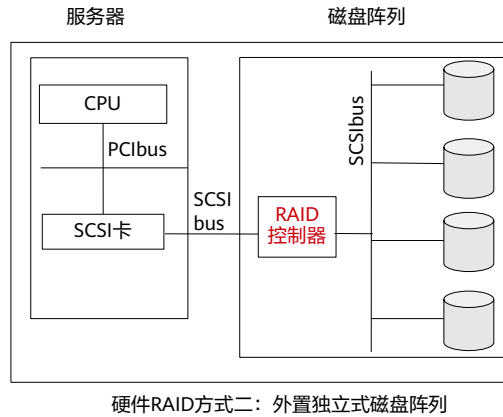
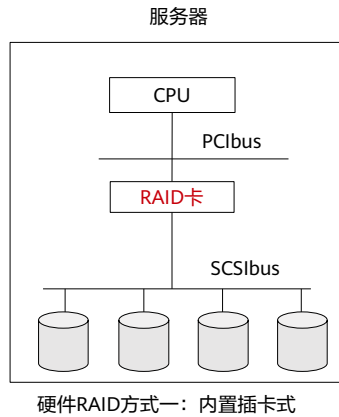
RAID热备、重构概念

- 热备（Hot Spare）的定义：
 - 当冗余的RAID组中某个硬盘失效时，在不干扰当前RAID系统的正常使用的情况下，用RAID系统中另外一个正常的备用硬盘自动顶替失效硬盘，及时保证RAID系统的冗余性。
- 热备一般分为两种：
 - 全局式：备用硬盘为系统中所有的冗余RAID组共享。
 - 专用式：备用硬盘为系统中某一组冗余RAID组专用。



- 数据校验，利用冗余数据进行数据错误检测和修复，冗余数据通常采用海明码、异或操作等算法来计算获得。利用校验功能，可以很大程度上提高磁盘阵列的可靠性、高性能和容错能力。不过，数据校验需要从多处读取数据并进行计算和对比，会影响系统性能。
- 一般来说，RAID不可作为数据备份的替代方案，它对非磁盘故障等造成的数据丢失无能为力，比如病毒、人为破坏、意外删除等情形。此时的数据丢失是相对操作系统、文件系统、卷管理器或者应用系统来说的，对于RAID而言，数据都是完好的，没有发生丢失。所以，数据备份、灾备等数据保护措施是非常必要的，与RAID相辅相成，保护数据在不同层次的安全性，防止发生数据丢失。

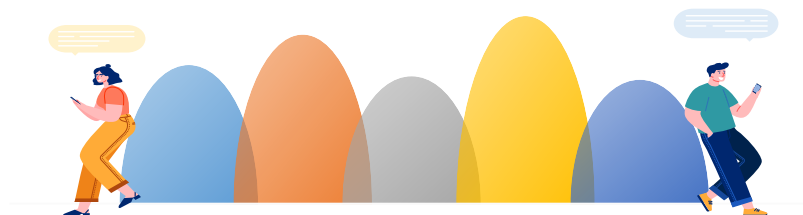
RAID的实现—硬件方式



- 基于硬件的RAID是利用硬件RAID适配卡来实现的。
- 硬件RAID又可分为内置插卡式和外置独立式磁盘阵列。
- RAID卡上集成了处理器，能够独立于主机对RAID存储子系统进行控制。因为拥有自己独立的处理器和存储器，RAID卡可以自己计算奇偶校验信息并完成文件定位，减少对主机CPU运算时间的占用，提高数据并行传输速度。

RAID的实现—软件方式

- 定义：
 - 软件RAID指包含在操作系统中，RAID功能完全用软件方式由系统的核心磁盘代码来实现。
- 特点：
 - 软件RAID不需要昂贵的RAID控制卡，提供了廉价的解决办法；
 - RAID功能完全依靠CPU执行，主机的CPU占用严重，如RAID5的大量异或（XOR）操作。



- 软件RAID中不能提供如下功能：
 - 硬盘热拔插；
 - 硬盘热备份；
 - 远程阵列管理；
 - 可引导阵列支持；
 - 在硬盘上实现阵列配置；
 - SMART硬盘支持。

RAID的实现—方式比较

方式	软件RAID	内置插卡式RAID	外置独立磁盘阵列RAID
特点	全部RAID功能靠CPU执行，主机CPU占用严重，系统性能下降。	减少了密集RAID操作对主CPU的占用，从而提高了性能。	通过标准控制器与服务器相连，RAID功能全部由外部RAID存储子系统上的微处理器来实现，独立于操作系统。
优点	• 实现成本低； • 配置灵活。	• 具有数据保护和高速特点； • 与软件RAID相比有更强的容错性和更好的性能； • 与外置相比价格相对低廉； • 支持可引导阵列。	• 为高端服务器提供超大容量存储系统； • 可配置双控制器，提高数据吞吐率，或供双机共享存储； • 支持热插拔； • 更好的可扩展性。

RAID卡

- RAID卡又名磁盘阵列卡，简称阵列卡，用以提高磁盘子系统的性能及可靠性。



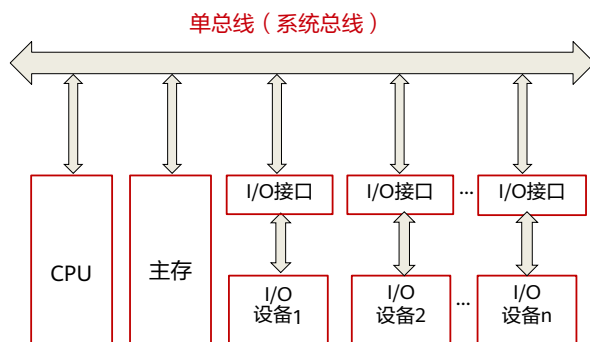
序号	名称
1	芯片过温指示灯
2	心跳灯
3	散热器
4	miniSAS接口 (port 0, SFF-8643)
5	miniSAS接口 (port 1, SFF-8643)

Avago MegaRAID SAS 9440-8i

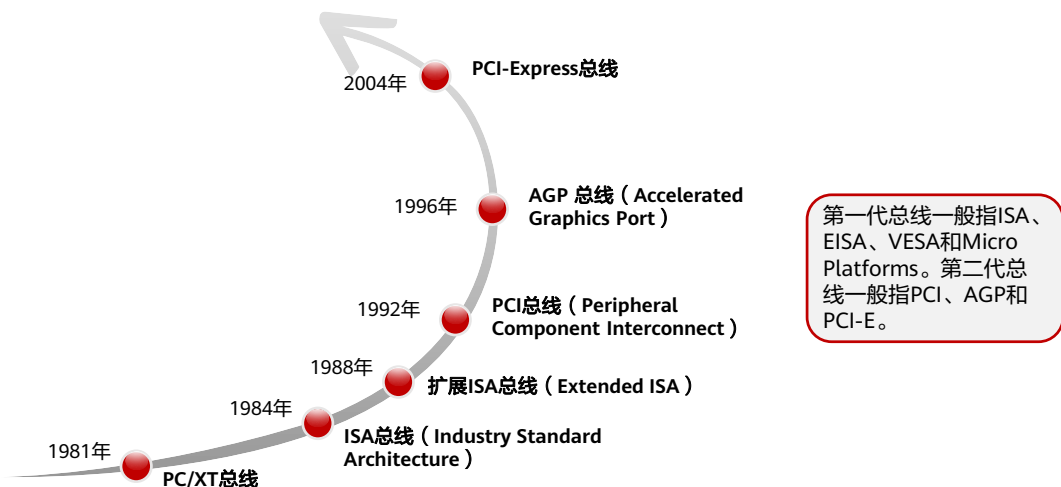
- RAID卡的作用：
 - 可以将若干硬盘驱动器按照一定要求组成一个整体、由阵列控制器管理的系统。
 - 可以提高磁盘子系统的性能及可靠性。

总线简介

- 总线（Bus）是计算机各种功能部件之间传送信息的公共通信干线。PCI-Express（peripheral component interconnect express），简称PCIe，是一种高速串行计算机扩展总线标准，主要用于扩充计算机系统总线数据吞吐量以及提高设备通信速度。



总线技术的发展史

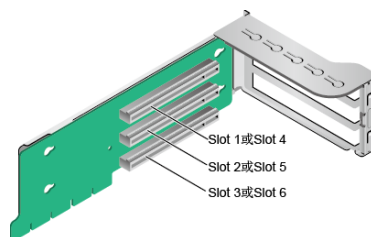
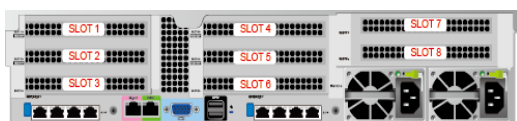


- 第一代总线一般指ISA、EISA、VESA和Micro Platforms。第二代总线一般指PCI、AGP和PCI-X
- ISA总线 (Industry Standard Architecture)
 - 最早的PC总线是IBM公司1981年在PC/XT电脑采用的系统总线，它基于8 bit的8088 处理器，被称为PC总线或者PC/XT总线。
 - 1984年，IBM推出基于16-bit Intel 80286处理器的PC/AT电脑，系统总线也相应地扩展为16 bit，并被称呼为PC/AT总线。而为了开发与IBM PC兼容的外围设备，行业内便逐渐确立了以IBM PC总线规范为基础的ISA（工业标准架构：Industry Standard Architecture）总线。
 - ISA总线最大传输速率仅为8 MB/s因此它的弱点也是显而易见的，传输速率过低、CPU占用率高、占用硬件中断资源等，很快使ISA总线在飞速发展的计算机技术中成为瓶颈。
 - 因此在1988年，康柏、惠普等9个厂商协同把ISA扩展到32-bit，这就是著名的EISA（Extended ISA）总线。EISA总线的工作频率仍旧仅有8 MHz，并且与8/16 bit 的ISA总线完全兼容，由于是32-bit 总线的缘故，带宽提高了一倍，达到了32 MB/s。可惜的是，EISA仍旧由于速度有限，并且成本过高，在还没成为标准总线之前，在20世纪90年代初的时候，就给PCI 总线给取代了。

- PCI总线（Peripheral Component Interconnect）
 - 由于ISA/EISA总线速度缓慢，一度出现CPU的速度甚至还高过总线的速度，造成硬盘、显示卡还有其它的外围设备只能通过慢速并且狭窄的瓶颈来发送和接受数据，使得整机的性能受到严重的影响。为了解决这个问题，1992年Intel 在发布486处理器的时候，也同时提出了32-bit的PCI（周边组件互连）总线。
 - 最早提出的PCI总线工作在33 MHz频率之下，传输带宽达到了133 MB/s（33 MHz X 32 bit/8），比ISA总线有了极大的改善，基本上满足了当时处理器的发展需要。目前计算机上广泛采用的是这种32-bit、33 MHz 的PCI 总线。
- AGP 总线（Accelerated Graphics Port）
 - PCI总线是独立于CPU的系统总线，可将显示卡、声卡、网卡、硬盘控制器等高速的外围设备直接挂在CPU总线上，打破了瓶颈，使得CPU的性能得到充分的发挥。可惜的是，由于PCI总线只有133 MB/s的带宽，对付声卡、网卡、视频卡等绝大多数输入/输出设备也许显得绰绰有余，但对于胃口越来越大的3D显卡却力不从心，并成为了制约显示子系统和整机性能的瓶颈。因此，PCI总线的补充——AGP 总线就应运而生了。
 - Intel 于1996年7月正式推出了AGP（加速图形接口，Accelerated Graphics Port）接口，这是显示卡专用的局部总线，是基于PCI 2.1版规范并进行扩充修改而成，工作频率为66 MHz，1X模式下带宽为266 MB/S，是PCI总线的两倍。后来依次又推出了AGP 2X、AGP4X，现在则是AGP 8X，传输速度达到了2.1 GB/S。
- PCI-Express
 - 随着新的技术和设备层出不穷，特别是游戏和多媒体应用越来越广泛，PCI 的工作频率和带宽都已经无法满足需求。此外，PCI还存在IRQ共享冲突，只能支持有限数量设备等问题。在经历了长达10年的修修补补，PCI总线已经无法满足电脑性能提升的要求，必须由带宽更大、适应性更广、发展潜力更深的新一代总线取而代之，这就是PCI-Express总线。2002年正式公布了PCI Express 1.0 规范。PCI Express2.0规范于2006年正式推出。相对于PCI总线来讲，PCI-Express总线能够提供极高的带宽，来满足系统的需求。

Riser卡和PCIe槽位

- Riser卡，是指插在PCI-E接口上的功能扩展卡或转接卡，它是新一代的总线接口。
- 以TaiShan 200服务器为例，结合本门课程前面介绍的后面板的模组及接口，IO模组1提供的槽位为Slot 1 ~ Slot 3；IO模组2提供的槽位为Slot 4 ~ Slot 6；IO模组3提供的槽位为Slot 7 ~ Slot 8。



PCIexpress总线技术

PCIe版本	推出时间	原始传输率	带宽				
			× 1	× 2	× 4	× 8	× 16
1	2003	2.5 GT/s	250 MB/s	0.50 GB/s	1.0 GB/s	2.0 GB/s	4.0 GB/s
2	2007	5.0 GT/s	500 MB/s	1.0 GB/s	2.0 GB/s	4.0 GB/s	8.0 GB/s
3	2010	8.0 GT/s	984.6 MB/s	1.97 GB/s	3.94 GB/s	7.88 GB/s	15.8 GB/s
4	2017	16.0 GT/s	1969 MB/s	3.94 GB/s	7.88 GB/s	15.75 GB/s	31.5 GB/s
5	2019	32.0 GT/s	3938 MB/s	7.88 GB/s	15.75 GB/s	31.51 GB/s	63.0 GB/s

传输通道数	脚Pin总数	主接口区Pin数	总长度	主接口区长度
x1	36	14	25 mm	7.65 mm
x4	64	42	39 mm	21.65 mm
x8	98	76	56 mm	38.65 mm
x16	164	142	89 mm	71.65 mm

网卡定义和功能

- 网卡定义

- 网卡，又称为网络适配器或网络接口卡NIC（Network Interface Card），是构成计算机网络系统中最基本的、最重要的、必不可少的连接设备，计算机主要通过网卡接入网络。
- 主芯片，是网卡的核心元件，一块网卡性能的好坏和功能的强弱多寡，主要取决于芯片的质量。

- 网卡主要功能

- 代表固定的网络地址
- 数据的发送与接收
- 数据的封装与解封
- 链路管理
- 编码与译码



网卡的基础知识

- 网卡的物理接口主要有两种：电口和光口。



TM210 灵活IO卡（4xGE电口）

电口：即普通的RJ45接口，连接网线。

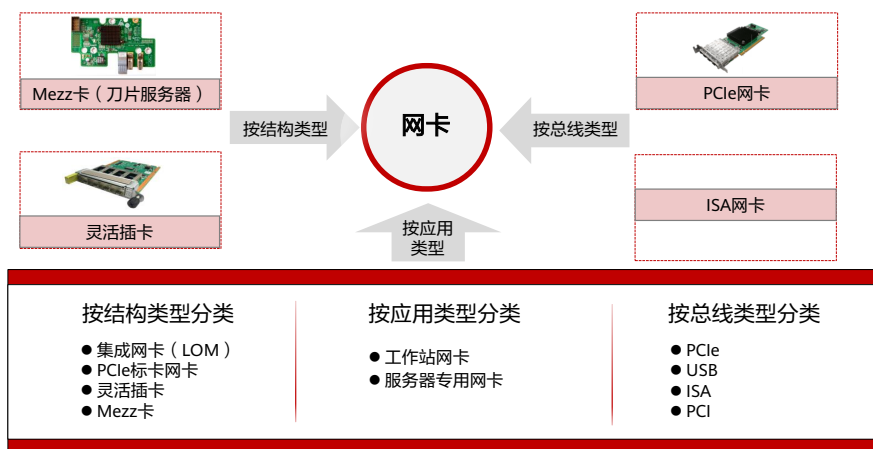


TM280 灵活IO卡（4x25 GE光口）

光口：用于连接光模块，根据接口封装形式，可以分为SFP+、SFP28、QSFP+。

- 网卡的物理接口有两种，一种是电口，一种是光口。
- 电口，也就是我们在普通PC上见到的那种网口接口，这种接口叫RJ45，使用的是普通的网线。
- 另一种接口是光口，用于连接光模块，网卡上用于插光模块的接口，我们叫光笼子。

网卡的分类



- 我们再来说说网卡的分类。随着计算机网络技术的飞速发展，为了满足各种环境和层次的应用，出现了不同类型的网卡。
- 总线分类：PCIe、USB、ISA、PCI，ISA/PCI这几种总线类型是比较早期的网络总线，现在已很少用了，USB接口的网卡主要用在消费级电子中。
- 结构形态：集成网卡（LOM）、PCIe标卡网卡、Mezz卡。
- 应用类型：按网卡所应用的计算机类型来区分，可以将网卡分为应用于工作站的网卡和应用于服务器的网卡。

华为服务器网卡

- 板载网卡
 - 集成在服务器主板的PCH芯片中，不可更换。
 - 对外提供2个GE电口+2个10 GE光口/电口，集成网卡不占用PCIe插槽。
- PCIe标卡
 - 华为有自研及外购的PCIe标卡网卡，可以配置在标准PCIe插槽中。
- 灵活IO插卡
 - 华为自研网卡，非标准PCIe卡形态，只能用于华为机架服务器。
- Mezz卡
 - 刀片服务器计算节点专用Mezz卡，只能用于华为E9000刀片服务器。



板载网卡



PCIe标卡



灵活IO插卡



Mezz卡

- 华为服务器的网卡，根据网卡的不同形态，可以分为：板载网卡，PCIe标卡网卡，灵活IO插卡，Mezz网卡。
 - 板载网卡，也就是我们常说的LOM网卡，是集成在服务器主板的PCH芯片中，不可更换的，不占用服务器的PCIe槽位；
 - PCIe标卡网卡，标卡指的是网卡的形态，为了解决网卡的结构兼容性，PCIe中定义了PCIe的尺寸，各网卡厂家按照这个尺寸开发的网卡，可以安装在PCIe标准的插槽中；
 - 灵活IO插卡，这个是华为机架服务器在设计上，为了更大的利用空间，开发了这种非标准结构的网卡，只能用于华为服务器上；
 - Mezz卡特指刀片的网卡，网卡不直接对外出物理接口，所有信号都是通过机框背板与外界相连。各个厂家的Mezz卡不能通用。

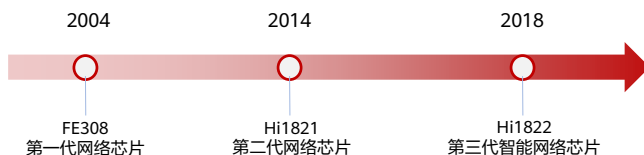
服务器专用网卡和普通网卡的差异对比

- 相对于服务器专用网卡来说，普通网卡指应用在普通PC、工作站、消费级电子产品中的网卡，对网卡的可靠性、安全性等要求不高；
- 服务器与普通PC的不同在于，服务器一直处于工作中，且要求长时间稳定运行。

区别	服务器网卡	普通网卡
应用场景	服务器	普通PC、工作站、消费电子产品中的网卡
传输速率	10 Gbps、25 Gbps	10 Mbps、100 Mbps
安全性	高，有容错功能	一般，无容错功能

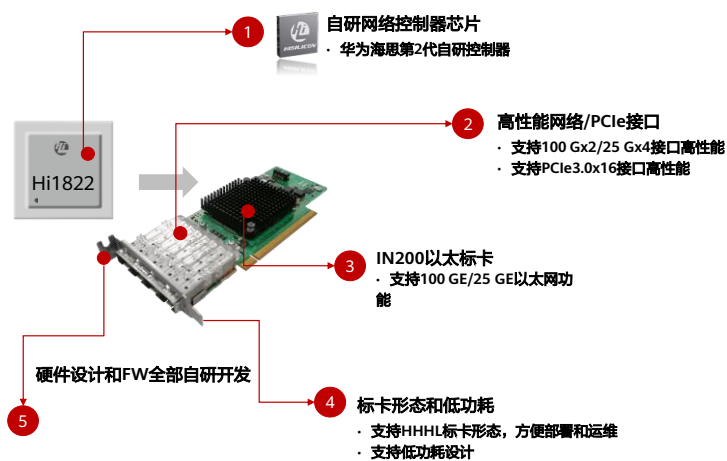
华为智能融合网络芯片

- 华为智能网卡主要应用于云网络OVS加速和云存储大规模组网低时延两大场景。



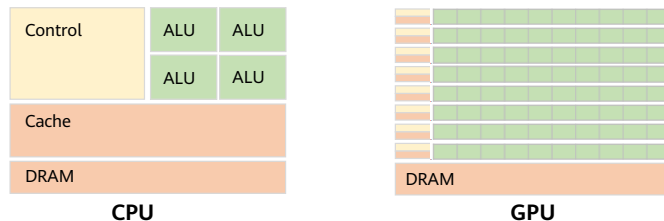
- 网络在生活中无处不在，像游戏视频、社交应用、金融交易、云支付等大吞吐量业务或低时延敏感业务对网络性能提出了苛刻的需求，网络加速已经成为业务发展的核心诉求。
- 从2004年华为就开始研发智能融合网络芯片，华为第三代智能融合网络芯片具体的特征有：以太与FC融合、25 GE~100 GE，16 G~32 G FC、内置48个可编程数据转发核心、OVS和RoCE v1/v2协议卸载、15 mpps OVS转发性能、支持SR-IOV等。

“IN” 智能网卡



GPU

- GPU英文全称Graphic Processing Unit，中文译为“图形处理器”。分为桌面级和企业级，服务器主要涉及企业级（工作站、服务器、高性能计算）。
- GPU与CPU的硬件逻辑架构对比：



- 其中Control是控制器、ALU算术逻辑单元、Cache是CPU内部缓存、DRAM就是内存。可以看到GPU设计者将更多的晶体管用作执行单元，而不是像CPU那样用作复杂的控制单元和缓存。从实际来看，CPU芯片空间的5%是ALU，而GPU空间的40%是ALU。这也是导致GPU计算能力超强的原因。

- GPU当前只是单纯的并行矩阵的乘法和加法运算，对于神经网络模型的构建和数据流的传递还是在CPU上进行。

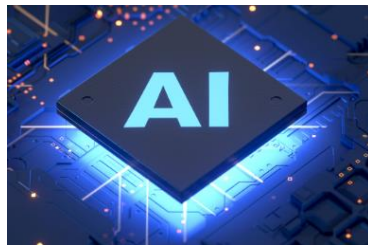
集成显卡和独立显卡

主要分类	类别	主流厂商、产品（及客户）
接入方式	独立GPU	AMD（Radeon系列）、NVIDIA（GeForce系列）
	集成GPU	Intel（HD系列）、AMD（APU系列）
应用端	PC GPU	intel、NVIDIA、AMD
	服务器GPU	NVIDIA（Tesla）、AMD（FireStream）
	移动GPU	ARM系列：华为麒麟处理器、三星猎户座、苹果A系列；高通骁龙

区别对比	集成显卡	独立显卡
与CPU的关系	集成在CPU里面的图像处理单元，是构成CPU的一部分	单独插在主板上的图像处理单元，其接口是PCIe，是一个独立的电脑组件
价格	低	高
兼容性	较好	较差
性能	较差	较好
功耗	低	高
是否占用电脑内存	是	否

NPU介绍

- NPU（Neural Networks Process Units，神经网络处理单元），工作原理是在电路层模拟人类神经元和突触，并且用深度学习指令集直接处理大规模的神经元和突触，一条指令完成一组神经元的处理。NPU通过突触权重实现存储和计算一体化，从而提高运行效率。
- NPU是模仿生物神经网络而构建的，CPU、GPU处理器需要用数千条指令完成的神经元处理，NPU只要一条或几条就能完成，因此在深度学习的处理效率方面优势明显。
- 与GPU一样，NPU同样需要CPU的协同处理才能完成特定的任务。

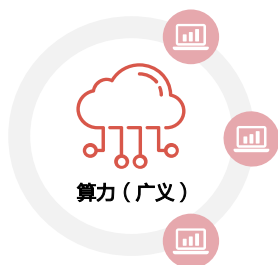


- NPU与GPU加速不同，主要体现为每层神经元计算结果不用输出到主内存，而是按照神经网络的连接传递到下层神经元继续计算，因此其在运算性能和功耗上都有很大的提升。
- CPU将编译好的神经网络模型文件和权重文件交由专用芯片加载，完成硬件编程。
- CPU在整个运行过程中，主要是实现数据的加载和业务流程的控制，其交互流程为：
 - 打开NPU专用芯片设备
 - 传入模型文件，得到模型task
 - 获取task的输入输出信息
 - 拷贝输入数据到模型内存中
 - 运行模型，得到输出数据

算力

- 算力即计算能力。算力的狭义定义是一台计算机具备的理论上最大的每秒浮点运算次数（FLOPS）。广义的算力，既包括了狭义的算力，也包括了存力和算法。
- 算力是数据中心的服务器通过对数据进行处理后实现结果输出的一种能力。

——《中国算力白皮书（2022年）》



算力（狭义）

CPU、GPU等芯片技术

存力

内存、硬盘等存储技术

算法

操作系统、数据库、中间件、应用程序等软件技术

如何衡量算力

- FLOPS (Floating-point operations per second, 每秒浮点运算次数), 通常被用于评估算力。浮点运算实际上包括了所有涉及小数的运算, 在某类应用软件中常常出现, 比整数运算更费时间。现今大部分的处理器中都有浮点运算器。因此每秒浮点运算次数所量测的实际上就是浮点运算器的执行速度。
 - 一个MFLOPS (megaFLOPS) 等于每秒一百万 ($=10^6$) 次的浮点运算;
 - 一个GFLOPS (gigaFLOPS) 等于每秒十亿 ($=10^9$) 次的浮点运算;
 - 一个TFLOPS (teraFLOPS) 等于每秒一万亿 ($=10^{12}$) 次的浮点运算;
 - 一个PFLOPS (petaFLOPS) 等于每秒一千万亿 ($=10^{15}$) 次的浮点运算;
 - 一个EFLOPS (exaFLOPS) 等于每秒一百亿亿 ($=10^{18}$) 次的浮点运算。
- 然而, 仅仅关注FLOPS并不足以全面评估处理器的性能。TOPS, 即teraOPS, 是另一个关键指标。它衡量的是处理器每秒可以执行多少万亿次操作。与TFLOPS不同的是, TOPS更多地关注于处理器的整数运算能力。



服务器的散热方式：风冷和液冷



风冷

- 主要由室内主机和室外冷凝器组成，效率高，易于冗余配置。
- 适应温度环境宽，维护简单。
- 有制冷铜管管路安装简便、可维护性强、系统可靠性高的优点，因此在中小型数据中心得到广泛应用。



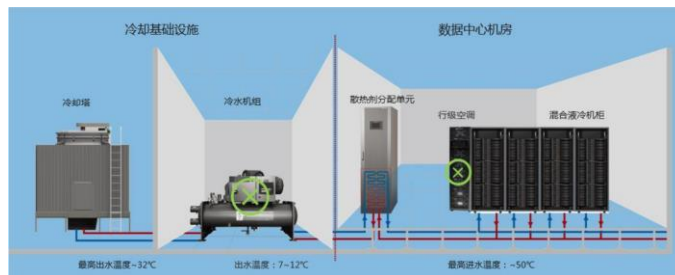
液冷

- 通过液体直接导向热源带走热量，不像风冷间接通过空气循环降温，效率更高、更节能且噪音更小。
- 液冷主要有冷板、浸没、喷淋三种部署方式。

- 风冷：受室内外机组安装距离限制、难以解决室内外机间超长冷媒管路的回油、阻力损失大，冷量衰减的问题。

液冷方案介绍

- 液冷方案主要分为直接式和间接式，直接式目前主要有：浸没（单相、两相-低沸点液体）、喷淋，间接式主要为冷板。
- 目前业界主要方案采用浸没和冷板居多，冷板一般只覆盖主要发热元器件，其他元器件热量需要采用其他方式散热。



57 Huawei Confidential



- 冷板（间接）
 - 液冷：将冷板固定在服务器的主要发热元器件上，液体流经冷板将热量带走。其它低散热元器件，则由液冷门或空调带走热量。
 - 安全度高：工作液体不与电子元器件直接接触，安全性高。
 - 易维护：冷却液采用去离子水（含一定比例乙二醇），安全无毒。闭式循环系统。
- 喷淋（直接）
 - 100%液冷：喷淋式系统是一个直接液冷系统，在服务器内部部署喷淋模块。根据发热体位置和发热量大小不同，让冷却液有针对性地对发热器件进行喷淋，达到设备降温的目的。
 - 效率高：喷淋的液体和被冷却器件直接接触，冷却效率高。
- 浸没（直接）
 - 100%液冷：将发热元件直接浸没在冷却液中，依靠液体的流动循环带走服务器等设备运行产生的热量。
 - 效率高：由于发热元件与冷却液直接接触，散热效率更高，噪音更低，可解决防高热谜底。
 - 单相/两相：浸没式液冷分为两相液冷和单相液冷，散热方式可以采用干冷器和冷却塔等形式。

服务器电源的定义及分类

- 服务器电源就是指使用在服务器上的电源（POWER），它和PC（个人电脑）电源一样，都是一种开关电源。
- 电源按照标准可以分为ATX电源和SSI电源两种。
 - **ATX标准**：输出功率一般在125瓦~350瓦之间；主要用于台式机、工作站和低端服务器。ATX电源通常采用20 Pin（20针）的双排长方形插座给主板供电。
 - **SSI标准**：SSI（Server System Infrastructure，服务器系统架构）规范是Intel联合一些主要的IA架构服务器生产商推出的新型服务器电源规范。SSI规范的推出是为了规范服务器电源技术，降低开发成本，延长服务器的使用寿命而制定的。



SSI电源

- SSI电源是一种电源解决方案，它可以为 ups和 hvdc提供电源。
- ATX标准是Intel在1997年推出的一个规范，输出功率一般在125 W~350 W之间。ATX电源通常采用20 Pin（20针）的双排长方形插座给主板供电。随着Intel推出Pentium4处理器，电源规范也由ATX修改为ATX12 V，和ATX电源相比，ATX12 V电源主要增加了一个4 Pin的12 V电源输出端，以便更好地满足P4的供电要求（2 GHz主频的P4功耗达到52.4 W）。
- SSI（ServerSystemInfrastructure）规范是Intel联合一些主要的IA架构服务器生产商推出的新型服务器电源规范。SSI规范的推出是为了规范服务器电源技术，降低开发成本，延长服务器的使用寿命而制定的，主要包括服务器电源规格、背板系统规格、服务器机箱系统规格和散热系统规格。
- IA架构是指智能化架构，是一种基于人工智能技术的软件架构。它主要包括数据采集、数据处理、模型训练和模型推理等环节，通过这些环节实现对数据的智能化处理和分析。IA架构可以应用于各种领域，如自然语言处理、计算机视觉、语音识别等。它可以帮助企业快速构建智能化应用，提高业务效率和竞争力。

服务器电源参数示例

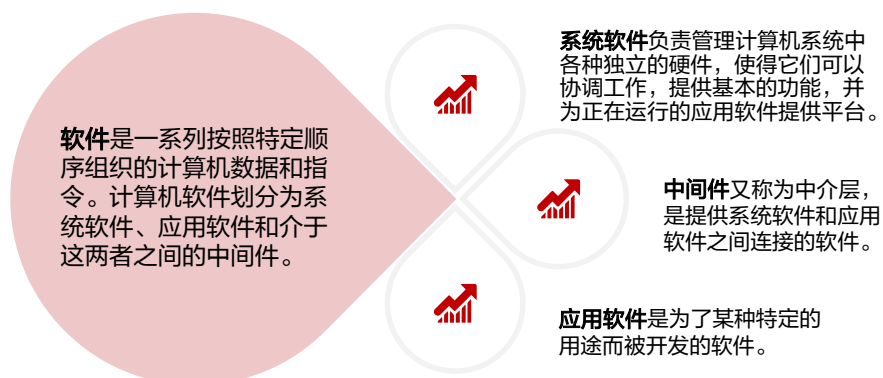
- 常见的服务器电源有73.5 mm和68.0 mm两种规格。
- 以TaiShan 200（型号2280）和2288X V5等机架式服务器为例，使用的是68mm尺寸的电源。

型号	CRPS 1300W服务器电源	CRPS 800W服务器电源	鲲鹏900W服务器电源	鲲鹏2000W服务器电源
电源外观				
输入电压	90 V AC~264V AC 180 V DC~320V DC	90 V AC~264 V AC 180 V DC~320 V DC	90 V AC~264 V AC 180 V DC~300 V DC	90 V AC~264 V AC 180 V DC~300 V DC
输出电压	11.4V DC~12.6V DC	12V DC&12Vsb DC	12 V DC	12 V DC
输出功率	12 V DC: 1300 W; 12 Vsb DC: 36 W	800 W	900 W	2000 W
尺寸（深*宽*高）	185.0 mm*73.5 mm*40.0 mm	185.0 mm*73.5 mm*40.0 mm	183.0 mm*68.0 mm*40.5 mm	183.0 mm*68.0 mm*40.5 mm

目录

1. 服务器硬件组成
- 2. 服务器软件介绍**
3. 操作系统介绍

软件基本概述



目录

1. 服务器硬件组成
- 2. 服务器软件介绍**
 - 服务器相关协议
 - BMC
 - BIOS
3. 操作系统介绍

IPMI

- IPMI定义

- IPMI（Intelligent Platform Management Interface，智能平台管理接口），是一种开放标准的硬件管理接口规格，定义了嵌入式管理子系统进行通信的特定方法。
- IPMI信息通过基板管理控制器（BMC）进行交流，使用低级硬件智能管理而不使用操作系统进行管理。



- IPMI是智能型平台管理接口（Intelligent Platform Management Interface）的缩写，是管理基于Intel结构的企业系统中所使用的外围设备采用的一种工业标准，该标准由英特尔、惠普、NEC、美国戴尔电脑和SuperMicro等公司制定。用户可以利用IPMI监视服务器的物理健康特征，如温度、电压、风扇工作状态、电源状态等。而且更为重要的是IPMI是一个开放的免费标准，用户无需为使用该标准而支付额外的费用。
- IPMI的发展：
 - 1998年Intel、DELL、HP及NEC共同提出IPMI规格，可以透过网路远端控制温度、电压。
 - 2001年IPMI从1.0版改版至1.5版，新增PCI Management Bus等功能。
 - 2004年Intel发表了IPMI 2.0的规格，能够向下相容IPMI 1.0及1.5的规格。新增了Console Redirection，并可以通过Port、Modem以及LAN远端管理服务器，也加强了安全、VLAN 和刀片服务器的支持性。

Redfish

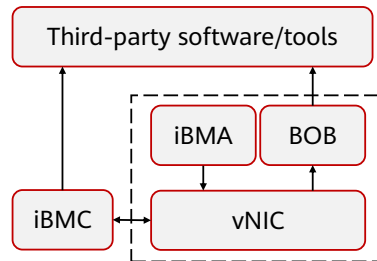
- Redfish是一种基于HTTPs服务的管理标准，利用RESTful接口实现设备管理。目前我们利用Redfish管理接口可以实现包括用户管理、获取服务器信息、管理模块信息等常用HDM和BIOS配置。当前由标准组织DMTF（www.dmtf.org）负责维护。
- 开发Redfish的主要原因之一是解决遗留IPMI无法有效解决的安全需求。



- 例如，Redfish基于其网络传输协议的超文本传输协议安全（HTTPS）/安全套接字层（SSL）标准，这比唯一的IPMI网络协议更普遍、更安全、更可审计。
- 在Redfish之前，现代数据中心环境缺少通用的管理标准。随着组织转向横向扩展解决方案，传统标准不足以满足管理众多单节点和多节点服务器融合架构的需要。
- Redfish从一开始就设计为融合基础架构的通用管理协议，提供了的强大的简便性，可以集成到现有的工具链中。

iBMA系统介绍

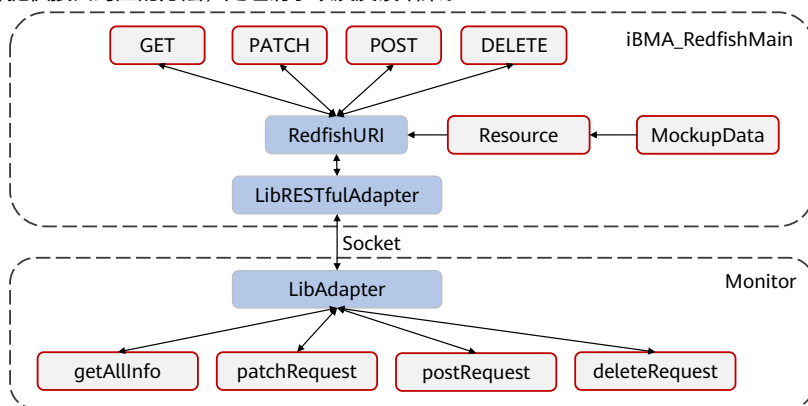
- iBMA 2.0作为带内管理代理软件，对上层管理软件iBMC开放了标准的Redfish管理接口，且可以通过BOB映射支持与带外系统的融合，以提供更完整的服务器设备管理能力。
- “带外+带内”管理系统模型：



- 带外和带内管理技术因为各自的优势，在服务器设备管理领域都得到了较广泛的应用。带外管理系统不对业务系统的运行带来影响，且在业务处理器故障情况下仍可正常工作；而带内管理系统可管理到一些带外无法获取的访问的信息，如进程信息、CPU性能状况、RAID控制器信息等。

Redfish接口

- Redfish服务通过注册路由的方式，提供接口，可以有GET、POST、PATCH、DELETE等操作；
- 底层插件提供接口对应的方法，处理请求以及反馈结果。



SNMP简介

- SNMP全称是：Simple Network Management Protocol，简单网络管理协议。
- SNMP是广泛应用于TCP/IP网络的网络管理标准协议，该协议能够支持网络管理系统，用以监测连接到网络上的设备是否有任何引起管理上关注的情况。
- SNMP采用轮询机制，提供最基本的功能集，适合小型、快速、低价格的环境使用，而且SNMP以用户数据报协议（UDP）报文为承载，因而受到绝大多数设备的支持，同时保证管理信息在任意两点传送，便于管理员在网络上的任何节点检索信息，进行故障排查。

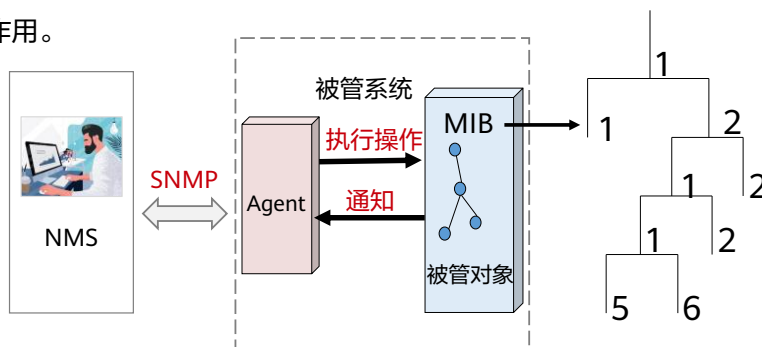


- SNMP是广泛应用于TCP/IP网络的网络管理标准协议，该协议能够支持网络管理系统，用以监测连接到网络上的设备是否有任何引起管理上关注的情况。SNMP采用轮询机制，提供最基本的功能集，适合小型、快速、低价格的环境使用，而且SNMP以用户数据报协议（UDP）报文为承载，因而受到绝大多数设备的支持，同时保证管理信息在任意两点传送，便于管理员在网络上的任何节点检索信息，进行故障排查。
- 为什么需要SNMP？
- 随着网络技术的飞速发展，在网络不断普及的同时也给网络管理带来了一些问题：网络设备数量成几何级数增加，使得网络管理员对设备的管理变得越来越困难；同时，网络作为一个复杂的分布式系统，其覆盖地域不断扩大，也使得对这些设备进行实时监控和故障排查变得极为困难。
- 网络设备种类繁多多样，不同设备厂商提供的管理接口（如命令行接口）各不相同，这使得网络管理变得愈发复杂。
- 在这种背景下，SNMP应运而生，SNMP是广泛应用于TCP/IP网络的网络管理标准协议，该协议能够支持网络管理系统，用以监测连接到网络上的设备是否有任何引起管理上关注的情况。

- 通过“利用网络管理网络”的方式：
- 网络管理员可以利用SNMP平台在网络上的任意节点完成信息查询、信息修改和故障排查等工作，工作效率得以提高。
- 屏蔽了设备间的物理差异，SNMP仅提供最基本的功能集，使得管理任务与被管理设备的物理特性、网络类型相互独立，因而可以实现对不同设备的统一管理，管理成本低。
- 设计简单、运行代价低，SNMP采用“尽可能简单”的设计思想，其在设备上添加的软件/硬件、报文的种类和报文的格式都力求简单，因而运行SNMP给设备造成的影响和代价都被最小化。

SNMP的基本组件

- SNMP基本组件包括NMS（Network Management System，网络管理系统）、Agent（代理进程）、被管对象（Managed Object）和MIB（Management Information Base，管理信息库）。如图所示它们共同构成SNMP的管理模型，在SNMP的体系结构中都起着至关重要的作用。

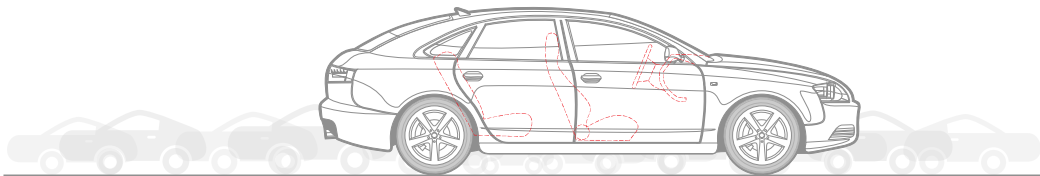


- NMS
- NMS在网络中扮演管理者角色，是一个采用SNMP协议对网络设备进行管理/监视的系统，运行在NMS服务器上。NMS可以向设备上的Agent发出请求，查询或修改一个或多个具体的参数值。
- NMS可以接收设备上的Agent主动发送的Trap信息，以获知被管理设备当前的状态。
- Agent
- Agent是被管理设备中的一个代理进程，用于维护被管理设备的信息数据并响应来自NMS的请求，把管理数据汇报给发送请求的NMS。Agent接收到NMS的请求信息后，通过MIB表完成相应指令后，并把操作结果响应给NMS。
- 当设备发生故障或者其它事件时，设备会通过Agent主动发送信息给NMS，向NMS报告设备当前的状态变化。
- Managed Object
- Managed Object指被管理对象。每一个设备可能包含多个被管理对象，被管理对象可以是设备中的某个硬件，也可以是在硬件、软件（如路由选择协议）上配置的参数集合。

- MIB
- MIB是一个数据库，指明了被管理设备所维护的变量，是能够被Agent查询和设置的信息。MIB在数据库中定义了被管理设备的一系列属性：对象的名称、对象的状态、对象的访问权限和对象的数据类型等。通过MIB，可以完成以下功能：Agent通过查询MIB，可以获知设备当前的状态信息。
- Agent通过修改MIB，可以设置设备的状态参数。
- SNMP的MIB采用树型结构，它的根在最上面，根没有名字。如下图所示的是MIB的一部分，它又称为对象命名树。每个对象标识符OID（object identifier）对应于树中的一个管理对象，该树的每个分支都有一个数字和一个名称，并且每个点都以从该树的顶部到该点的完整路径命名，如system的OID为1.3.6.1.2.1.1，interfaces的OID为1.3.6.1.2.1.2。

SNMP如何工作？

- 一旦网络中启动SNMP协议，NMS作为整个网络的网管中心，会对设备进行管理。
- 每个被管理设备都包含驻留在设备上的Agent、多个被管对象和MIB，NMS通过与运行在被管理设备上的Agent交互，由Agent通过对设备端的MIB的操作，完成NMS的指令。
- SNMP的工作原理是将协议数据单元（也称为SNMP GET请求）发送到响应SNMP的网络设备。
- 用户通过网络监控工具可以跟踪所有通信过程，并从SNMP获取数据。



目录

1. 服务器硬件组成
- 2. 服务器软件介绍**
 - 服务器相关协议
 - BMC
 - BIOS
3. 操作系统介绍

BMC介绍

- BMC定义

- BMC: Baseboard Management Controller, 主板管理控制单元, 是IPMI规范的核心, 负责各路传感器的信号采集、处理、储存, 以及各种器件运行状态的监控。BMC向机框管理板提供被管理对象的硬件状态及告警等信息, 从而实现机框管理模块对被管理对象的设备管理功能。

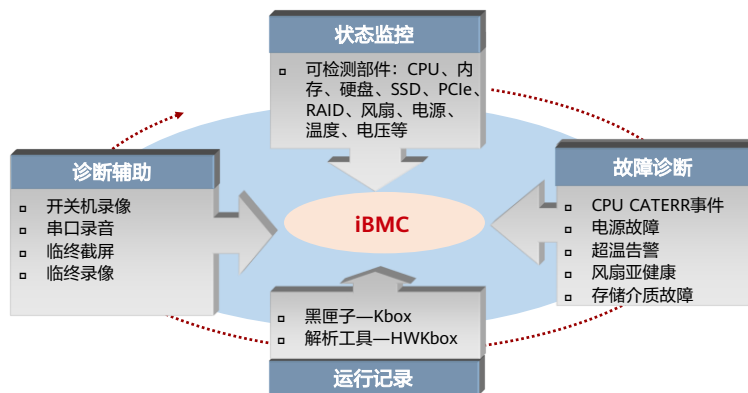


- BMC主要实现以下功能:

- 远程控制;
- 告警管理;
- 状态检测;
- 设备信息管理;
- 散热控制;
- 支持ipmitool工具;
- 支持Web界面管理;
- 支持集中账号管理。

iBMC介绍

- iBMC (Huawei Intelligent Baseboard Management Controller, 华为服务器智能管理系统) 是具有完全自主知识产权, 面向服务器全生命周期的服务器嵌入式管理系统。



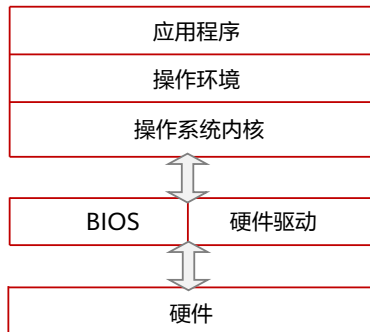
- iBMC提供硬件状态监控、部署、节能、安全等系列管理工具, 标准化接口构建服务器管理更加完善的生态系统。iBMC基于华为自研的管理芯片Hi1710, 采用多项创新技术, 全面实现服务器的精细化管理。
- iBMC提供了丰富的用户接口, 如命令行、基于Web界面的用户接口、IPMI集成接口、SNMP集成接口、Redfish集成接口, 并且所有用户接口都采用了认证机制和高度安全的加密算法, 保证接入和传输的安全性。

目录

1. 服务器硬件组成
- 2. 服务器软件介绍**
 - 服务器相关协议
 - BMC
 - BIOS
3. 操作系统介绍

BIOS简介

- BIOS：Basic Input/Output System，基本输入输出系统。
- BIOS实际上是固化到计算机中的一组程序，为计算机系统提供最底层的、最直接的硬件控制服务，解决硬件的实时需求。



BIOS的功能：

- 检测和初始化硬件
- 操作系统引导
- 高级电源管理

- BIOS是系统内核和硬件层之间桥梁。
- BIOS特点：
 - 软件升级、加载和装载功能；
 - 基本OAM功能；
 - 串口管理功能；
 - 故障恢复功能；
 - ECC管理功能；
 - 硬件诊断功能。

UEFI定义和特点

- 定义

- UEFI (Unified Extensible Firmware Interface , 统一可扩展固件接口) 是一种个人电脑系统规格, 用来定义操作系统与系统固件之间的软件界面, 作为BIOS的替代方案。可扩展固件接口负责加电自检 (POST)、联系操作系统以及提供连接操作系统与硬件的接口。

—维基百科

- 特点

- UEFI提供更大的磁盘容量
- UEFI提供更高的效能
- 64位新系统有优势
- 更方便批量安装
- 更快开机、休眠恢复
- 更安全启动



传统BIOS和UEFI运行流程

- 传统BIOS运行流程



- UEFI运行流程

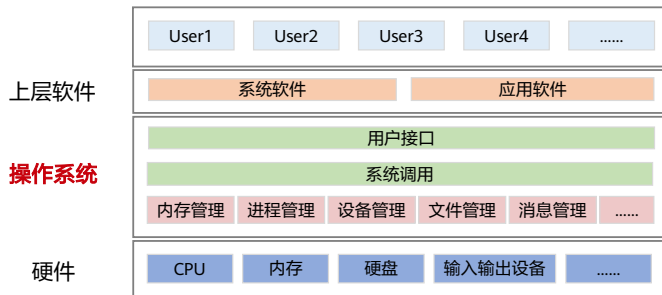


目录

1. 服务器硬件组成
2. 服务器软件介绍
- 3. 操作系统介绍**
 - 操作系统简介
 - 当前主流操作系统

操作系统简介

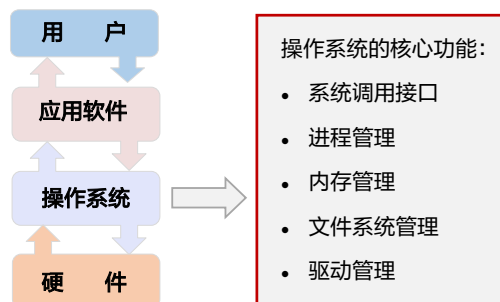
- OS（Operating System，操作系统）是一组主管并控制计算机操作、运用和运行硬件、软件资源和提供公共服务来组织用户交互的相互关联的系统软件程序，同时也是计算机系统的内核与基石。
- 操作系统需要处理如管理与配置内存、决定系统资源供需的优先次序、控制输入与输出设备、操作网络与管理文件系统等基本事务。操作系统也提供一个让用户与系统交互的操作界面。
- 操作系统由操作系统的内核（运行于内核态，管理硬件资源）以及系统调用（运行于用户态，为应用程序员写的应用程序提供系统调用接口）两部分组成。



- 操作系统是一种控制计算机的特殊计算机程序，在计算机和用户之间起到连接的作用，常常用来在多用户使用的情况下协调分配有限资源。这些资源包括CPU、磁盘、内存和打印机等，用户运行程序时需要访问这些资源。
- 主流操作系统
 - 根据应用领域来划分，可分为：
 - 桌面操作系统、服务器操作系统、主机操作系统、嵌入式操作系统。
 - 根据源码开放程度，可分为：
 - 开源操作系统（Linux、Unix）、闭源操作系统（Windows、Mac OS）。

操作系统的作用

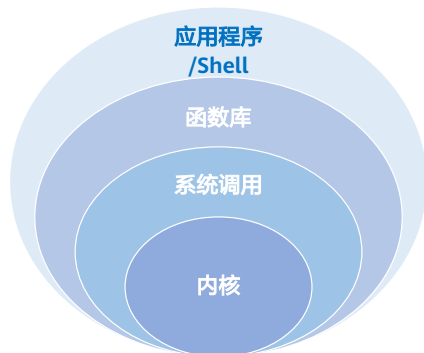
- 隐藏了硬件调用接口，为应用程序员提供调用硬件资源的更好，更简单，更清晰的模型（系统调用接口）。为应用程序提供如何使用硬件资源的抽象；
- 管理硬件资源，将应用程序对硬件资源的竞态请求变得有序化。现代的操作系统运行同时运行多道程序，操作系统的任务是在相互竞争的程序之间有序地控制对处理器、存储器以及其他I/O接口设备的分配。



- 现代的计算机系统主要是由一个或者多个处理器，主存，硬盘，键盘，鼠标，显示器，打印机，网络接口及其他输入输出设备组成。一般而言，现代计算机系统是一个复杂的系统。管理这些硬件并且加以优化使用是非常繁琐的工作，这个繁琐的工作就是操作系统来干的，应用软件直接使用操作系统提供的功能来间接使用硬件。
- 功能一：隐藏了硬件调用接口，为应用程序员提供调用硬件资源的更好，更简单，更清晰的模型（系统调用接口）。应用程序员有了这些接口后，就不用再考虑操作硬件的细节，专心开发自己的应用程序即可。
- 例如：操作系统提供了文件这个抽象概念，对文件的操作就是对磁盘的操作，有了文件我们无需再去考虑关于磁盘的读写控制（比如控制磁盘转动，移动磁头读写数据等细节）。
- 功能二：将应用程序对硬件资源的竞态请求变得有序化。
- 例如：很多应用软件其实是共享一套计算机硬件，比方说有可能有三个应用程序同时需要申请打印机来输出内容，那么a程序竞争到了打印机资源就打印，然后可能是b竞争到打印机资源，也可能是c，这就导致了无序，打印机可能打印一段a的内容然后又去打印c...，操作系统的一个功能就是将这种无序变得有序。

操作系统的组成

- 通常站在用户角度，操作系统由内核和各种应用程序组成。即系统分为用户空间和内核空间。
- 用户空间即上层应用程序活动的空间。
- 内核本质上是一种软件程序，用来管理计算机硬件资源并提供上层应用程序运行所需要的系统调用接口。



- 系统调用：应用程序的执行必须依托于内核提供的资源，包括CPU、存储、I/O资源等，为了使上层应用能够访问到这些资源，内核必须为上层应用提供访问接口—系统调用。（如果写一个汉字代表一个上层应用，那么一个系统调用更像是汉字的一个笔画）。
- 库函数：通过库函数实现对系统调用的封装，将简单的业务逻辑接口呈现给用户，方便用户调用。对于简单的操作我们可以直接调用系统调用来访问资源，对于复杂操作，可以借助库函数来实现。（库函数更像是汉字的偏旁）。
- Shell：是一个特殊的应用程序（俗称命令行），本质上是一个命令解释器，它可以执行符合Shell语法的文本（脚本），实际上这些Shell语句通常就是对系统调用做了一层封装，方便客户使用。

操作系统类别

- 从使用场景来看，操作系统大致可以分为桌面操作系统、服务器操作系统、移动终端操作系统、嵌入式操作系统、物联网操作系统、云操作系统等类别。

操作系统类别	代表
服务器操作系统	Unix (AIX、HPUX...) 、Linux、Windows Server ...
桌面操作系统 (PC、笔记本电脑等)	Windows (XP/10/11) 、macOS、Linux ...
移动终端操作系统 (手机、可穿戴设备等)	Android、iOS、Windows Phone (WP) 、HarmonyOS ...
嵌入式操作系统	FreeRTOS、Unison、eCos、NXPMQX ...
物联网操作系统	Android Things、ALiOS、ARM Mbed OS、LiteOS ...
云操作系统	Windows、CentOS、Ubuntu、EulerOS ...

目录

1. 服务器硬件组成
2. 服务器软件介绍
- 3. 操作系统介绍**
 - 操作系统简介
 - 当前主流操作系统

常用服务器操作系统分类

服务器操作系统	Windows Server	UNIX	Linux
优点	▫ 极其易用 ▫ 极大降低使用者的学习成本	▫ 支持大型文件系统服务、数据服务应用 ▫ 功能强大 ▫ 稳定性和安全性能好	▫ 开源系统 ▫ 非常成熟 ▫ 拥有完整的权限机制 ▫ 安全性与稳定性都很高
缺点	▫ 对服务器硬件要求较高 ▫ 稳定性不是很好	主要以命令的方式进行，不容易掌握	系统操作需要一定时间的学习

常用服务器操作系统

UNIX

一种多用户、多进程的计算机操作系统。支持大型文件系统服务、数据服务应用，功能强大、稳定性和安全性好。

常见UNIX OS:

HP-UX、IBM-AIX、Solaris、Apple UNIX。

GNU/Linux

Linux是类Unix计算机操作系统的统称。Linux系统比较成熟，拥有着一套完整的权限机制，安全性与稳定性都很高。

常见Linux OS:

SUSE、Kylin、RedFlag、CentOS、Redhat、openEuler。

Windows

Windows Server是微软发布的服务器操作系统，主要应用于服务器场景，拥有友好的用户视窗操作界面。

常见Windows Server版本:

2000、2003、2008、2012、2016、2019。

- Linux与Unix的关系:

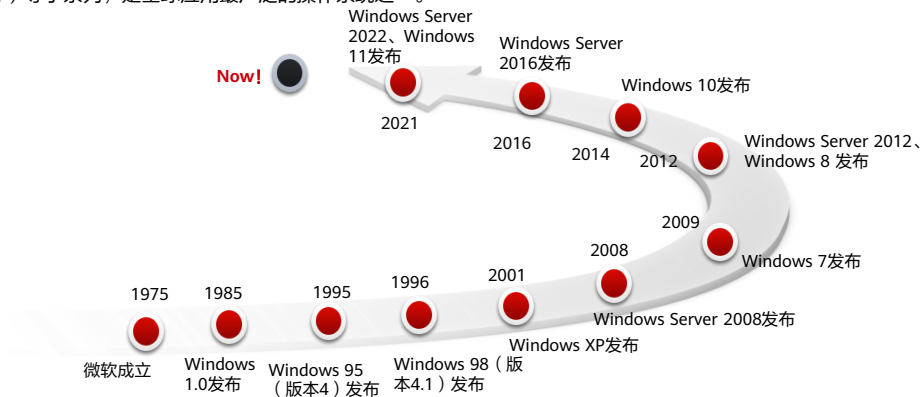
- Linux是一个类似UNIX的操作系统，并在功能和用户体验上进行优化，所以Linux模仿了UNIX，使得Linux在外观和交互上与UNIX非常类似。
- Linux内核最初是由李纳斯·托瓦兹（Linus Torvalds）在赫尔辛基大学读书时出于个人爱好而编写的，当时他觉得教学用的迷你版UNIX操作系统Minix太难用了，于是决定自己开发一个操作系统。第1个版本于1991年9月发布，当时仅有10000行代码。
- UNIX系统大多是与硬件配套的，也就是说，大多数UNIX系统如AIX、HP-UX等是无法安装在x86服务器和个人计算机上的，而Linux则可以运行在多种硬件平台上。
- UNIX是商业软件，而Linux是开源软件，是免费、公开源代码的。

- GNU计划:

- GNU计划，又称革奴计划，是由Richard Stallman在1983年9月27日公开发起的。它的目标是创建一套完全自由的操作系统。
- GNU是“GNU's Not UNIX”的递归缩写。Linux提供了一个操作系统内核，而GNU提供了大量的自由软件来丰富在其之上的各种应用程序。

Windows的发展

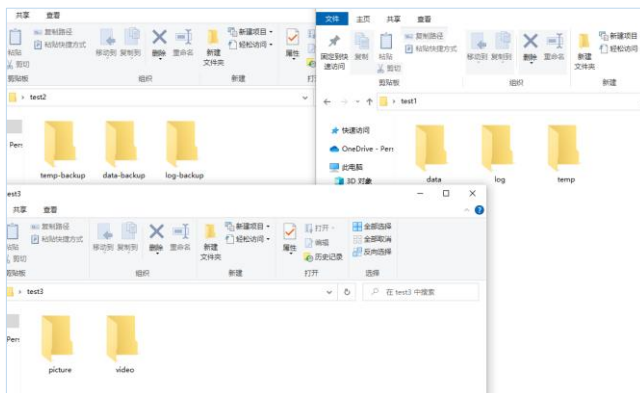
- Microsoft Windows是美国微软公司以图形用户界面为基础研发的操作系统，主要运用于计算机、智能手机等设备。共有普通版本、服务器版本（Windows Server）、手机版本（Windows Phone）、嵌入式版本（Windows CE、Windows for IoT）等子系列，是全球应用最广泛的操作系统之一。



- 1995年8月24日，微软推出具有里程碑意义的Windows 95。Windows 95是单用户、多任务操作系统，它能够在同一个时间片中处理多个任务，充分利用了CPU的资源空间，并提高了应用程序的响应能力。
- Windows NT发布后，开始了其在操作系统的领导角色。
- 2013年10月18日发布Windows Server 2012 R2。

Windows操作简单方便

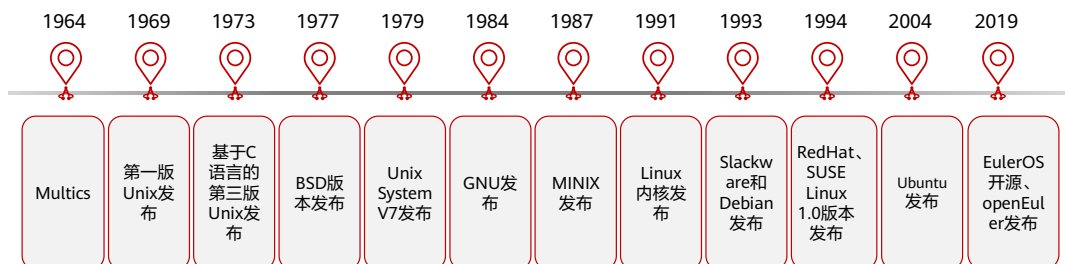
- Windows操作系统能够作为个人计算机的主流操作系统，其优异的人机操作性是重要因素。Windows操作系统界面友好，窗口制作优美，操作动作易学，多代系统之间有良好的传承，计算机资源管理效率较高，效果较好。



- Windows操作系统因操作界面友好，相较于Linux更加大众化，广泛应用于PC、管理Server等用户侧场景。

Unix/Linux的发展

- Linux是一种自由和开放源码的类UNIX操作系统。该操作系统的内核由李纳斯·托瓦兹在1991年10月5日首次发布，再加上用户空间的应用程序之后，就成为了Linux操作系统。Linux最初是作为支持英特尔x86架构的个人电脑的一个自由操作系统。目前Linux已经被移植到更多的计算机硬件平台，远远超出其他任何操作系统。



Linux操作系统特性

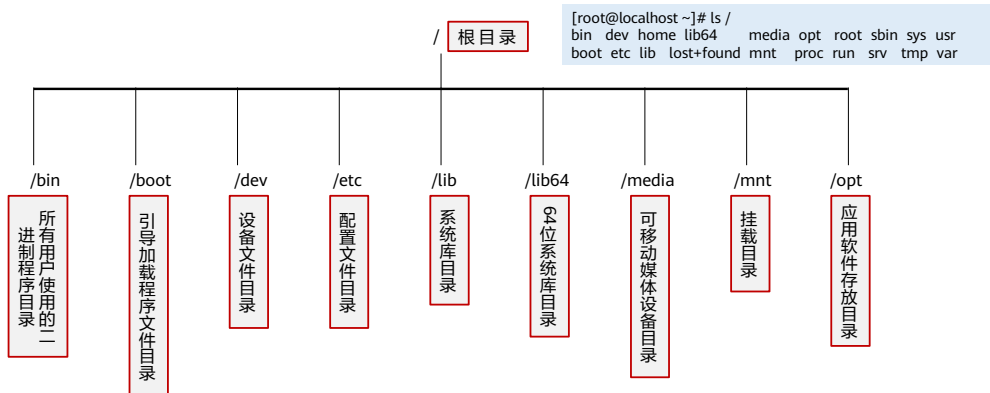
- 支持多种平台
 - Linux可运行在多种硬件平台上。此外，Linux还是一种嵌入式操作系统，可运行在掌上电脑、机顶盒等设备。
- 多用户多任务
 - 系统资源可以被不同用户各自拥有使用，同时执行多个程序，而且各个程序的运行相互独立。
- 自由软件
 - 用户可以随时免费获取它的源代码，并且可以根据自己的需求对源代码进行编辑、修改。
- 完全兼容POSIX1.0标准
- 继承了Unix的设计理念
 - 一切皆文件。

- POSIX，可移植操作系统接口（Portable Operating System Interface，缩写为POSIX），是IEEE为要在各种UNIX操作系统上运行的软件，而定义API的一系列互相关联的标准的总称，其正式称呼为IEEE 1003，而国际标准名称为ISO/IEC 9945。它基本上是Portable Operating System Interface（可移植操作系统接口）的缩写，而X则表明其对Unix API的传承。

- Linux操作系统是一种非常流行的多任务、多用户操作系统，应用非常广泛。Linux的主要特点为：
 - 多任务（Multi-tasking）：Linux是一个多任务操作系统，在它内部允许有多个任务同时运行。而DOS操作系统是单任务的操作系统，不能同时运行多个任务。系统在执行多任务时，CPU在某一时刻只能执行一个任务，但Linux操作系统将CPU时间分片，并把这些时间片分别安排给多个进程。因为CPU运行很快，在操作者看来，所有程序（进程）都在同时运行。
 - 多用户（Multi-users）：Linux又是一个多用户操作系统，它允许多个用户同时使用。在Linux中，每位用户运行自己的或公用的程序，好象拥有一台单独的机器。DOS操作系统是单用户的操作系统，只允许一个用户同时使用。
 - 管道：Linux允许一个程序的输出作为另外一个程序输入，多个程序串起来好像一条管道一样。通过各个简单任务的组合，就可以完成比较复杂的任务，并极大提高了操作的方便性。晚期版本的DOS操作系统也借鉴并提供了这种机制。
 - 功能强大的Shell：Linux的命令解释器由Shell实现。Linux提供了多种功能强大的Shell，每种Shell本身就是一种解释型高级语言，通过用户编程就可创造无数命令，使用方便。
 - 安全保护机制：Linux提供了非常强大的安全保护机制，防止系统及其数据未经许可而被非法访问。

Linux文件目录结构

- 在Linux操作系统中，一切皆是文件。
- Linux的文件目录结构是树形，“/”称为根目录。



- Linux下的哲学核心思想是“一切皆文件”。“一切皆文件”，指的是对所有文件（目录、字符设备、块设备、套接字、打印机、进程、线程、管道等）操作，读写都可用`fopen()` / `fclose()` / `fwrite()` / `fread()` 等函数进行处理。
- 登录系统后，在当前命令窗口下输入命令：`ls /` 你会看到上图所示的回显，那么这些目录是什么意思呢？以下是这些目录的解释：
- **/bin**：bin是Binary的缩写, 这个目录存放着最经常使用的命令。
- **/boot**：这里存放的是启动Linux时使用的一些核心文件，包括一些连接文件以及镜像文件。
- **/dev**：dev是Device（设备）的缩写, 该目录下存放的是Linux的外部设备，在Linux中访问设备的方式和访问文件的方式是相同的。
- **/etc**：这个目录用来存放所有的系统管理所需要的配置文件和子目录。
- **/lib**：这个目录里存放着系统最基本的动态连接共享库，其作用类似于Windows里的DLL文件。几乎所有的应用程序都需要用到这些共享库。
- **/media**：Linux系统会自动识别一些设备，例如U盘、光驱等等，当识别后，Linux会把识别的设备挂载到这个目录下。
- **/mnt**：系统提供该目录是为了让用户临时挂载别的文件系统的，我们可以将光驱挂载在/mnt/上，然后进入该目录就可以查看光驱里的内容了。
- **/opt**：这是给主机额外安装软件所摆放的目录。比如你安装一个ORACLE数据库则就可以放到这个目录下。默认是空的。

思考题

1. （多选题）以下关于华为服务器网卡的描述，正确的是哪些项？
 - A. 板载网卡集成在服务器主板的PCH芯片中，不可更换
 - B. 华为自研的PCIe标卡网卡，可以配置在标准PCIe插槽中
 - C. 灵活IO插卡集成在服务器面板上，可用于前端业务连接
 - D. Mezz卡可用于华为机架服务器
2. （判断题）主板管理控制单元（BMC）是IPMI规范的核心，负责各路传感器的信号采集、处理、储存，以及各种器件运行状态的监控。
 - A. 正确
 - B. 错误

- 参考答案：
- AB
- A

本章总结

- 本章我们一起学习了服务器的硬件、软件架构和组成
- 了解了服务器CPU、内存、硬盘、RAID卡等硬件原理
- 了解了操作系统的工作原理，并熟悉了常见的操作系统

学习推荐

- 鲲鹏社区：
 - <https://www.hikunpeng.com/>
- 华为云官方链接：
 - <https://www.huaweicloud.com/>

缩略语

- GPU: Graphics Processing Unit, 图形处理器, 又称显示核心、视觉处理器、显示芯片, 是一种专门在个人电脑、工作站、游戏机等设备上图像运算工作的微处理器, 可用于桌面的图片渲染。
- NPU: Neural Processing Unit, 神经处理单元, 是一类专用于人工智能 (特别是人工神经网络、机器视觉、机器学习等) 硬件加速的微处理器或计算系统。
- CPU: Central Processing Unit, 中央处理器是计算机的主要设备之一, 其功能是解释计算机指令以及处理计算机软件中的数据, 与内部存储器、输入及输出设备成为现代电脑的三大部件。
- CISC: Complex Instruction Set Computer, 复杂指令集计算机, 是一种微处理器指令集架构, 每个指令可执行若干低端操作, 诸如从存储器读取、存储、和计算操作, 全部集于单一指令之中。
- RISC: Reduced Instruction Set Computer, 精简指令集计算机, 或简译为精简指令集, 是计算机中央处理器的一种设计模式。这种设计思路可以想像成是一家模块化的组装工厂, 对指令数目和寻址方式都做了精简, 使其实现更容易, 指令并行执行程度更好, 编译器的效率更高。

缩略语

- PC: Program Counter, 程序计数器, 是一个中央处理器CPU中的寄存器, 其功能是存储下一条要执行的指令 (在机器语言中) 的内存地址。
- IR: Instruction Register, 指令寄存器, 在计算机科学中是中央处理器中控制单元用来存储执行中指令的寄存器。
- ID: Instruction Decoder, 指令译码器, 是控制器中的主要部件之一。计算机能且只能执行“指令”。指令由操作码和地址码组成。操作码表示要执行的操作性质, 即执行什么操作, 或做什么; 地址码是操作码执行时的操作对象的地址。
- OC: Operation Controller, 操作控制器, 是CPU控制器组成部分之一, 用来产生各种操作控制信号。
- SPD: Serial Presence Detect, 串行存在检测, 是1个8针的SOIC封装 (3mm*4mm) 256字节的EEPROM (Electrically Erasable Programmable ROM 电可擦写可编程只读存储器) 芯片。型号多为24LC01B, 位置一般处在内存条正面的右侧, 里面记录了诸如内存的速度、容量、电压与行、列地址带宽等参数信息。

缩略语

- RAM: Random Access Memory, 随机存储器, 基于半导体的可被CPU或者其他硬件设备读写的内存, 可以任何顺序访问存储位置。
- ROM: Read Only Memory, 只读存储器, 以非破坏性读出方式工作, 只能读出无法写入信息。信息一旦写入后就固定下来, 即使切断电源, 信息也不会丢失, 所以又称为固定存储器。
- DIMM: Dual Inline Memory Module, 双列直插内存模块, 一种新型的内存条, 提供了64位的数据通道, 有168条引脚。
- SDRAM: synchronous dynamic random-access memory, 同步动态随机存取存储器, 是有一个同步接口的动态随机存取内存 (DRAM)。
- DWPD: Drive Write Per Day, 每日整盘写入次数。
- NVMe: Non-Volatile Memory Express, 非易失性存储器标准, 是一种用于高度并行数据传输的协议, 可减少闪存存储和固态硬盘 (SSD) 中使用的每个输入/输出 (I/O) 的系统开销。
- RAID: Redundant Array of Independent Disks, 独立硬盘冗余阵列。

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。
Bring digital to every person, home, and
organization for a fully connected,
intelligent world.

Copyright©2024 Huawei Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.



华为计算产业介绍



前言

- 通过前面章节的学习，我们了解了通用服务器的组件和操作系统基础知识。本章我们将一起来学习鲲鹏一系列处理器和计算整机、鲲鹏生态体系如openEuler和openGauss等和鲲鹏HPC解决方案。
- 同时，我们还会深入学习昇腾的一系列硬软件产品如处理器、CANN、MindSpore等。了解昇腾行业应用和华为人工智能计算中心。

目标

- 学完本课程后，您将能够：
 - 了解什么是鲲鹏，熟悉鲲鹏芯片和整机产品；
 - 掌握鲲鹏生态体系内容，熟悉openEuler和openGauss的特点及优势；
 - 了解什么是昇腾，熟悉昇腾硬件架构及其硬件产品；
 - 了解昇腾AI计算平台CANN和其他昇腾软件产品；
 - 了解昇腾行业应用和华为人工智能计算中心。

目录

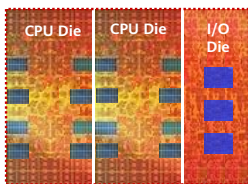
1. 鲲鹏计算产业概述

- 什么是鲲鹏
 - 鲲鹏处理器
 - 鲲鹏计算整机
 - 鲲鹏生态
 - 鲲鹏HPC解决方案

2. 昇腾计算产业概述

鲲鹏简介

鲲鹏是SoC



高性能：具备业界领先的56G高速SerDes能力，主板性能提升25%

高可靠：信号误码率低于10的负12次方，故障率比业界平均水平低15%

高能效：创新的DEMT动态节能技术，可以实现能效比领先业界15%以上

鲲鹏是计算平台

处理器->单机->集群，鲲鹏开放硬件平台



鲲鹏开放主板



鲲鹏计算整机

完备的软件工具链，发挥鲲鹏最佳性能



鲲鹏开发套件
DevKit



鲲鹏应用使能套件
BoostKit

鲲鹏是生态应用

使能合作伙伴

- 应用
- 中间件
- 数据库
- 操作系统
- 服务器/PC

使能行业应用

- 大数据
- 分布式存储
- 高性能计算
- 原生应用
- 云服务

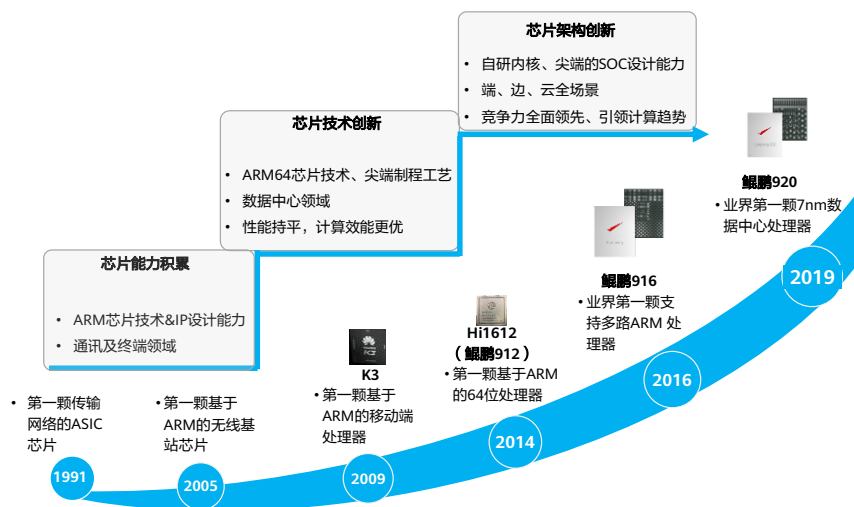
目录

1. 鲲鹏计算产业概述

- 什么是鲲鹏
- 鲲鹏处理器
- 鲲鹏计算整机
- 鲲鹏生态
- 鲲鹏HPC解决方案

2. 昇腾计算产业概述

鲲鹏处理器—坚持持续创新



- 华为在芯片及处理器领域的探索已经超过了二十年。
- 1991年，研发了第一颗用于传输网络的ASIC芯片。
- 2005年，研发了第一颗基于ARM架构的无线基站芯片。
- 2009年，推出智能手机CPU K3，麒麟芯片的前身，如今麒麟芯片已经把手机带入智慧时代。
- 2016年，推出第一代鲲鹏处理器，面向数据中心应用的鲲鹏916处理器。推出服务器CPU，也就是鲲鹏芯片的前身。

鲲鹏处理器简介

- 鲲鹏处理器是基于ARM指令集开发的通用处理器，是华为自主研发的基于ARM架构的企业级系列处理器产品，包含“算、存、传、管、智”五个产品系统体系。
- 物理架构上包含SoC、Chip、DIE、cluster、core等概念。



- SoC: System on Chip, 称为系统级芯片，也有称片上系统，意指它是一个产品，是一个有专用目标的集成电路，其中包含完整系统并有嵌入软件的全部内容。例如，Kunpeng 920除了CPU外，还集成了RoCE 网卡、SAS控制器和南桥。
- Chip: 芯片，是指有大规模集成电路的硅片，我们见过的CPU这种是最常见的芯片。一般几块硅片可以封装在一起组成一个芯片。
- DIE: 芯片的最小物理单元，Kunpeng 920封装了3个DIE，两个用来做计算，第三个用来做IO。
- Cluster: 若干个核(core)的集合。Kunpeng 920把4个core集合成为一个cluster,而一个DIE上有8个cluster。
- Core: 真正的计算单元，我们在操作系统侧看到的“核”。

华为鲲鹏处理器架构（ARM）特点

- 华为鲲鹏处理器基于ARM架构。有别于Intel、AMD CPU采用的CISC复杂指令集，ARM CPU采用RISC精简指令集。
- 优点：
 - 采用ARM架构，同样功能性能占用的芯片面积小、功耗低、集成度更高，更多的硬件CPU核具备更好的并发性能。
 - 支持64位指令集，能很好地兼容从IOT、终端到云端的各类应用场景。
 - 大量使用寄存器，大多数数据操作都在寄存器中完成，指令执行速度更快。
 - 采用RISC指令集，指令长度固定，寻址方式灵活简单，执行效率高。
- 不足：
 - 在数据中心领域属于新进入者，其生态仍处于快速发展阶段。

- 芯片工艺进入28 nm以下空间时，靠先进工艺提升性能成本急剧上升，工艺、主频遇到瓶颈后，开始转向增加核数的横向扩展来提升性能。
- 一个ARM核的面积仅为x86核的1/7，同样的芯片尺寸下，ARM的核数是x86的4倍以上。

华为鲲鹏通用计算处理器

- 华为从2004年开始基于ARM技术自研芯片，在通用计算处理器领域，2014年发布Kunpeng 912处理器，2016年发布Kunpeng 916处理器，2019年1月发布Kunpeng 920处理器，Kunpeng 920处理器是业界第一颗采用7 nm工艺的数据中心级的ARM架构处理器。



**鲲鹏912
(Hi1612)**

第一颗基于ARM
的64位CPU

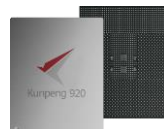
2014



鲲鹏916

业界第一颗支持
多路ARM CPU

2016



鲲鹏920

业界第一颗7 nm数据中
心 ARM CPU

2019

鲲鹏处理器特点



多核架构

2个CPU Die，每个CPU Die包括最多32个核、4个DDR4内存控制器

1个IO Die，提供PCIe接口、以太网接口、存储控制器、片间缓存一致接口和硬件加速引擎等功能

多合一 SoC

CPU | 南桥 | 网卡 | 硬盘控制器等多个芯片功能合一，支持多种数据加解密、压缩/解压缩机制

片间互联

华为Cache一致性总线（HCCS），为内核、设备、集群提供系统内存的一致访问

片间带宽最高可达480 Gbps，实现最多4个鲲鹏920处理器互联和最高256个物理核的NUMA架构

华为鲲鹏处理器极致性能



高性能

性能: 2.9x ↑

高吞吐

内存带宽: 60% ↑
I/O 总带宽: 66% ↑
网络带宽: 4x

高集成

1 颗 = 4颗芯片

- 单处理器整型计算性能，相比上一代提升2.9倍，提升大数据、分布式存储和数据库等场景的并行计算性能。
- 支持8个DDR4内存通道，内存带宽提升60%。
- 支持PCIe 4.0，IO带宽提升66%。
- 支持100GE以太网网络，网络带宽提升4倍。
- 支持同时搭载4颗鲲鹏芯片。

华为鲲鹏处理器技术创新



内核全自研

- 自研CPU内核，相比上一代提升2.9倍
- 自研片间支持2路/4路互联
- 自研多种硬件协加速引擎



内存/网络接口&IO协议

- 支持8通道DDR4内存控制器
- 支持100G RoCE端口
- 支持PCIe 4.0/CCIX协议



制程工艺领先

- 业界首款7纳米数据中心ARM CPU
- 采用业界领先的CoWoS封装技术，实现多Die合封



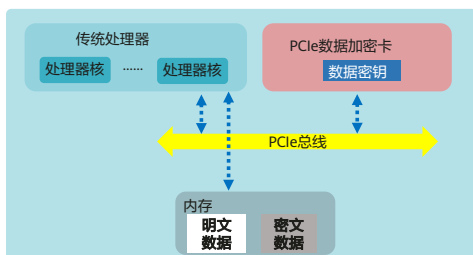
可靠性提升

- 支持使用概率高、收益大的RAS特性45条，实现在RAS特性上的增强

华为鲲鹏加速引擎增强数据安全

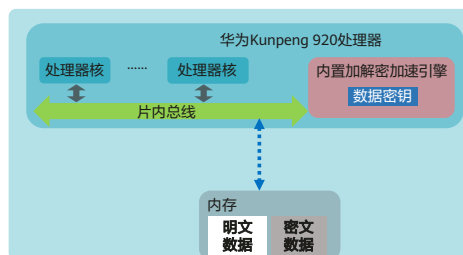
传统PCIe加密卡方案

- 明文数据通过PCIe总线传输，有数据泄密风险



华为Kunpeng安全加解密方案

- 华为Kunpeng内置加解密加速引擎，不占用计算资源
- 明文数据仅通过片内总线传输，安全性高
- 支持SM3/SM4国密算法加速



- 华为Kunpeng 920处理器数据的加解密在CPU内执行，密钥存储在安全区，只有特定进程能写入密钥，只能CPU直接读入到CPU，加强了数据的安全性。

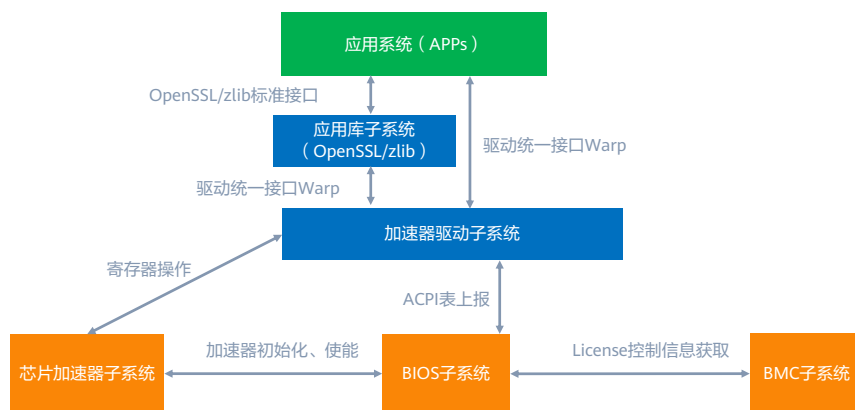
华为鲲鹏加速器应用场景

- 鲲鹏KAE利用内置的硬件加速单元和特有的指令，大幅度提升应用性能，使用特有框架，帮助降低调用路径性能损耗。
- 拥有了鲲鹏KAE的加速能力，相当于系统的每个CPU额外集成了性能强大的压缩或加解密硬件加速卡。

场景	引擎	数据
大数据	KAE	流数据
数据加密	KAE	块数据
智能安防	KAE	视频流数据
Web 服务	KAE	握手连接

- 加速引擎是TaiShan 200服务器基于华为Kunpeng 920提供的硬件加速解决方案，包含了对称加密、非对称加密和数字签名，压缩解压缩等算法，用于加速SSL/TLS应用和数据压缩，可以显著降低处理消耗，提高处理器效率。此外，加速引擎对应用层屏蔽了其内部实现细节，用户通过 OpenSSL、zlib标准接口即可实现快速迁移现有业务。
- KAE：Kunpeng Acceleration Engine，即华为 Kunpeng加速器引擎。

华为鲲鹏加速器系统逻辑架构



- 华为Kunpeng 920加速器系统逻辑架构：
 - 芯片加速器子系统、BIOS子系统和BMC子系统为TaiShan硬件产品自带子系统：
 - 芯片加速器子系统：集成在华为Kunpeng 920中，提供加速器的能力，对上层提供寄存器接口；
 - BIOS子系统：单板BIOS软件系统，主要负责根据License决定对加速器哪些模块初始化，并上报加速器ACPI表到内核（加速器驱动子系统处理）；
 - BMC子系统：服务器BMC软件系统，在这里主要负责对加速器License的管理。
 - 加速器驱动子系统：向上层提供各子加速器模块统一的驱动接口，是本系统的核心子系统；
 - 应用库子系统：应用库子系统包括OpenSSL加速器引擎、zlib替代库等，向上层提供标准接口；
 - 应用系统：用户系统，通过调用应用库子系统或驱动子系统实现加速器的功能。
- 目前加速引擎主要支持以下算法：
 - 摘要算法SM3；
 - 对称加密算法SM4，支持CTR/CBC模式；
 - 非对称算法RSA，支持异步模型，支持Key Sizes 1024/2048/3072/4096；

- 压缩解压缩算法，支持zlib/gzip。

目录

1. 鲲鹏计算产业概述

- 什么是鲲鹏
- 鲲鹏处理器
- 鲲鹏计算整机
- 鲲鹏生态
- 鲲鹏HPC解决方案

2. 昇腾计算产业概述

基于华为鲲鹏处理器，构建整机计算能力

高效能计算

提供兼容ARM架构的高性能
华为鲲鹏处理器、TaiShan
服务器和解决方案，将高效
能计算带入数据中心；

安全可靠

华为鲲鹏处理器基于自研内
核，TaiShan服务器计算芯片
全自研；
17年计算创新铸就稳如泰山
的高品质；

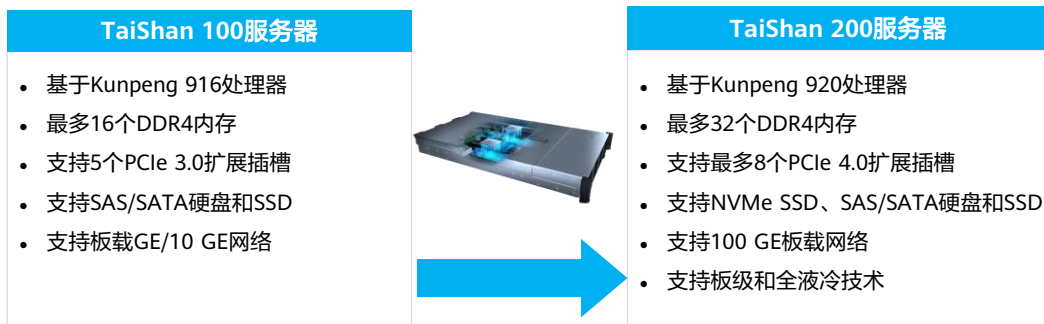
开放生态

开放平台，支持业界主流软
硬件；构建鲲鹏生态，与开
发者、伙伴和产业组织共同
打造智能计算新底座；



- TaiShan服务器以华为鲲鹏处理器为基础，构建整机计算能力。
- 第一代TaiShan 100服务器是基于鲲鹏916处理器，2016年推出市场。2019年推出TaiShan 200服务器基于最新的鲲鹏920处理器，是市场的主打产品。

TaiShan服务器系列介绍



- 第一代TaiShan 100服务器是基于鲲鹏916处理器，2016年推出市场。2019年推出TaiShan 200服务器基于最新的鲲鹏920处理器，是市场的主打产品。

TaiShan 200服务器全景图



- 基于Kunpeng处理器，华为打造了鲲鹏计算整机产品标杆—TaiShan服务器。
- TaiShan服务器是过往17年华为在计算技术和整机工程方面长期积累的结晶。
- 首先，华为在TaiShan 200服务器上充分应用了散热液冷、高速互联、可靠性设计与质量品控等工程工艺技术，将TaiShan服务器打造成为精品，为Kunpeng处理器应用在数据中心服务器领域树立了一个行业标杆。
- TaiShan服务器目前已经规模商用的有2280均衡型、5280存储型。
- 除此之外还包括支持72盘位的5290高密度型，支持4路服务器互联的2480高性能型，1U双路的1280高密度型和适合在边缘计算场景部署的2280E型号。

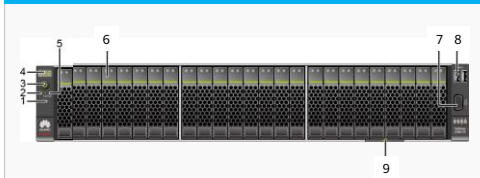
2280均衡型规格与亮点

2280 均衡型	前视图	型号	2280	主要特点
		形态	2U2路	高性能 - 鲲鹏920处理器，性能比肩x86高端型号 8通道内存技术，支持32个DDR4内存插槽，最高内存容量可达4 TB - 支持华为Atlas 300 AI加速卡、ES3000 V5 NVMe SSD
		CPU	2*鲲鹏920处理器，可支持32，48，64核主频2.6 GHz，最高功耗180 W	
		内存	32个DDR4 DIMM，最高2933 MT/s，最大内存4 TB	
		本地存储	16*3.5"SAS/SATA HDD硬盘 27*2.5"AS/SATA HDD硬盘 16*2.5"NVMe SSD硬盘	灵活适配 - 支持多个IO模组，实现丰富的硬盘配置 - 支持板载的灵活网卡，支持GE/10 GE/25 GE，实现不同网络配置
		RAID支持	RAID0/1/5/6/10/50/60	
		PCIe扩展	最多8个PCIe 4.0 x8或3个PCIe 4.0 x16 + 2个PCIe 4.0 x8标准插槽	安全可靠 - 采用华为全自研计算芯片 - 整机器件实现全国产化，保障可持续供应
		板载网卡	4*GE OR 4*10 GE OR 4*25 GE	
		电源	2个热插拔900 W交流或2000 W交流或1200 W直流或1500 W直流电源模块，支持1+1冗余	
		风扇	4个热插拔风扇，支持N+1冗余	
		温度范围	5°C~40°C	

- 8通道内存技术，比x86 V5 CPU多2个通道，因此可以支持32个DDR4内存插槽，最高内存容量可达4TB。

2280型号前后面板

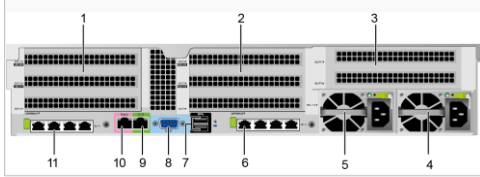
前面板



指示及按钮

按钮	指示	按钮	指示
1	UID按钮/指示灯	2	健康状态指示灯
3	电源开关按钮/指示灯	4	故障诊断数码管
5	灵活IO卡在位指示灯（1，2）	6	硬盘—
7	VGA	8	USB 3.0
9	标签卡		

后面板

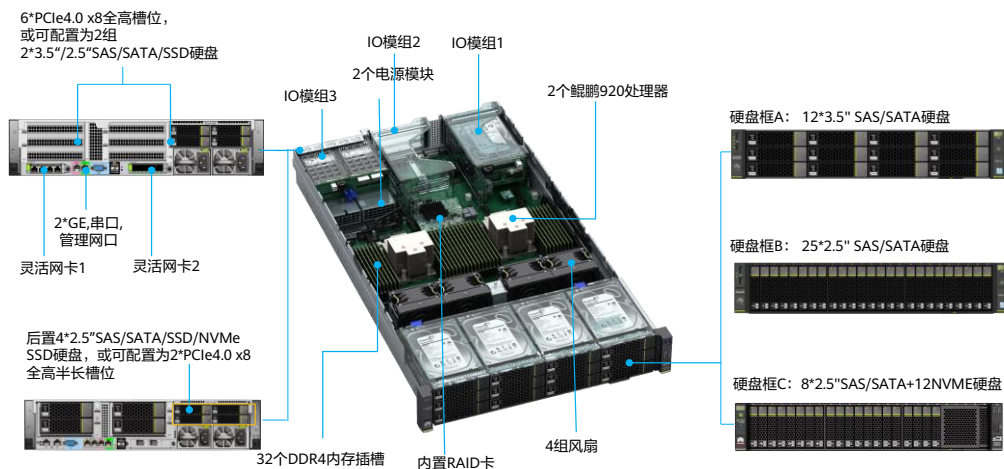


指示及按钮

按钮	指示	按钮	指示
1	IO模组1	2	IO模组2
3	IO模组3	4	电源模块2
5	电源模块1	6	灵活IO卡2
7	USB 3.0接口	8	VGA接口
9	调试串口	10	管理网口
11	灵活IO卡1	—	—

- 后面板的IO模组1和IO模组2可以选配基本上一致的硬盘模组或者Riser模组，IO模组3则不同。

2280型号内部结构



- 跟x86服务器最大的不同是CPU是焊接在主板上不能更换的，因此目前有3款不同的主板。另外内存条的密度也比x86大。

2280型号机型配置

类别	前置	后置	RAID方案
1 12盘位硬RAID方案	 12*3.5"SAS/SATA硬盘	 4*3.5"SAS/SATA硬盘 4*2.5"NVMe硬盘	RAID卡
2 12盘位SAS直通方案	 12*3.5"SAS/SATA硬盘	 2*2.5"SAS/SATA硬盘 4*2.5"NVMe硬盘	CPU SAS直通+软RAID
3 25盘位通用硬盘方案	 25*2.5"SAS/SATA硬盘	 2*2.5"SAS/SATA/硬盘 4*2.5"NVMe硬盘	RAID卡
4 25盘位NVMe SSD方案	 8*2.5"SAS+12*2.5"NVMe硬盘	 4*2.5"NVMe硬盘	RAID卡

- 2280有4种不同配置。

5280存储型规格与亮点

5280 存储型	前视图	型号	5280	主要特点
		形态	4U2路	超大存储 - 支持最多40个3.5"硬盘，本地存储容量可达560 TB - 鲲鹏920处理器，性能比肩x86高端型号，高效支持数据存储 8通道内存技术，支持32个DDR4内存插槽，最高内存容量可达4 TB 灵活适配 - 支持多个IO模组，实现丰富的硬盘配置 - 支持板载的灵活网卡，支持GE/10 GE/25 GE，实现不同网络配置 安全可靠 - 采用华为全自研计算芯片 - 整机器件实现全国产化，保障持续供应
		CPU	2*鲲鹏920处理器，可支持32，48，64核，2.6 GHz频率可选；最高功耗200 W	
		内存	32个DDR4 DIMM插槽，最高2933 MT/s	
		本地存储	40*3.5"SAS/SATA HDD或SSD硬盘 4*2.5"NVMe SSD硬盘	
		RAID支持	RAID0/1/5/6/10/50/60	
		PCIe扩展	最多8个PCIe 4.0 x8或3个PCIe 4.0 x16 + 2个PCIe 4.0 x8标准插槽	
		灵活网卡	4*GE OR 4*10 GE OR 4*25 GE	
		电源	2个热插拔2000 W交流或1200 W直流或1500 W直流电源模块，支持1+1冗余	
		风扇	4个热插拔风扇，支持N+1冗余	
		温度范围	5°C~35°C	

- 4U跟2U的机型是共主板的，只是硬板背板不同，支持最多40个3.5" 硬盘。
- 前面板24块，后面板12块及3个IO模组。

5280型号前后面板

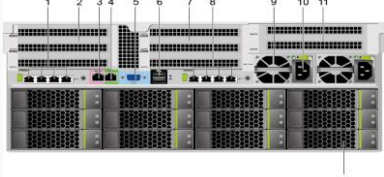
前面板



指示及按钮

按钮	指示	按钮	指示
1	UID按钮/指示灯	2	健康状态指示灯
3	电源开关按钮/指示灯	4	故障诊断数码管
5	灵活IO卡在位指示灯（1，2）	6	硬盘
7	VGA	8	USB 3.0
9	标签卡		

后面板

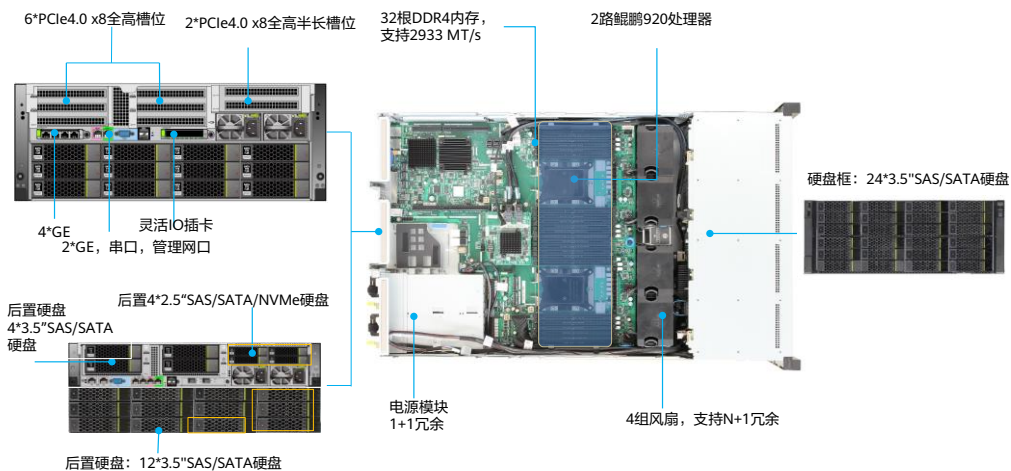


指示及按钮

按钮	指示	按钮	指示
1	灵活IO卡1	2	IO模组1
3	管理网口	4	调试串口
5	VGA接口	6	USB 3.0接口
7	IO模组2	8	灵活IO卡2
9	电源模块	10	电源模块接口
11	IO模组3	12	硬盘

- 5280前后面板跟2280基本一致，只是前后硬盘背板差异因此面板有所不同。

5280型号内部结构



- 5280内部结构跟2280大致相同，最主要的差别是多了下面12块硬盘的连接。

TaiShan机架服务器特点



2280均衡型



5280存储型

高性能

超强算力：高性能Kunpeng 920处理器。

大内存容量：8通道内存技术，支持32个DDR4内存插槽。

超强AI加速：支持华为Atlas 300 AI加速卡，满足多场景推理。

灵活适配

灵活的IO扩展：支持灵活插卡和标准的智能网卡，实现丰富的网络配置。

分级存储：支持大容量存储硬盘和ES3000 V5 NVMe SSD。

安全可靠

安全、可供应：采用华为全自研计算芯片，整机器件全国产化。

可靠与质量保障：减振、散热等高可靠设计；质量流程与测试。

- 3大特点：高性能、网络 and 存储配置灵活、安全可靠。

鲲鹏计算产业的典型应用场景

大数据



- ◆ 性能提升30%
- ◆ 加解密引擎
- ◆ 与x86融合部署

分布式存储



- ◆ IOPS性能提升20%
- ◆ 压缩解压引擎
- ◆ 与x86融合部署

数据库



- ◆ RoCE低时延网络
- ◆ 多核调度算法
- ◆ NUMA优化算法

原生应用



- ◆ 原生同构，无性能损耗
- ◆ 企业级可靠方案
- ◆ 百万级应用，成熟生态

云服务



- ◆ 鲲鹏裸金属服务器
- ◆ 鲲鹏弹性云服务器
- ◆ 鲲鹏Kubernetes容器

高性能

多核高并发



高吞吐

原生算力同构

- 鲲鹏计算产业的发展壮大，离不开产业界上下游伙伴的共同努力。在过去几年时间里，华为与各行业ISV共同孵化、完善了五大鲲鹏计算解决方案，在多个行业实现了规模商用。
- 譬如大数据场景，鲲鹏920处理器最高支持64核，多核优势意味着可并行处理更多的数据。实测数据显示，鲲鹏处理器的大数据性能比传统平台提升30%以上。
- 分布式存储场景，鲲鹏处理器通过内置数据压缩引擎，多核并行计算提升分布式存储软件的运行效率，相比传统处理器算法方案减少66%的压缩时间，并实现IOPS性能提升20%以上。
- 在鲲鹏处理器设计之初，华为就充分考虑了如何满足分布式并行计算的需求。在鲲鹏天生多核的基础上，增强了高I/O带宽、高内存带宽、高数据吞吐的设计，更好的满足了IT分布式演进的发展趋势。

目录

1. 鲲鹏计算产业概述

- 什么是鲲鹏
- 鲲鹏处理器
- 鲲鹏计算整机
- 鲲鹏生态
- 鲲鹏HPC解决方案

2. 昇腾计算产业概述

鲲鹏计算产业生态策略

- 华为作为鲲鹏计算产业的成员，基于“硬件开放、软件开源、使能伙伴、发展人才”的策略推动鲲鹏计算产业的发展。

合作伙伴 华为



使能合作伙伴

加速行业应用创新

提供开发套件与应用使能套件，支持伙伴快速迁移应用，使能极简开发、极致性能

软件开源

使能伙伴商业发行

- openGauss企业级开源数据库，使能伙伴数据库商业发行
- openEuler数字基础设施开源操作系统，使能伙伴操作系统商业发行

硬件开放

坚持伙伴优先

- 提供基于鲲鹏主板，SSD/网卡/RAID卡等部件，使能伙伴发展自有品牌部件/服务器和PC等

华为聚焦

为世界提供最强算力

- 计算架构创新，全场景芯片，云服务

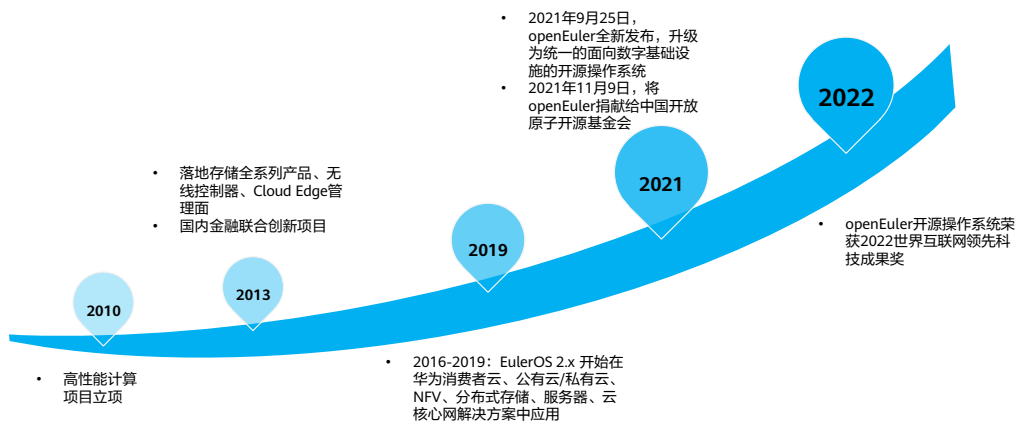
openEuler概述

- openEuler脱胎于EulerOS，EulerOS是华为公司自2010年起研发使用的服务器操作系统，Linux发行版之一，名字来源于著名数学家莱昂哈德·欧拉（Leonhard Euler）。
- 2019年9月，EulerOS正式开源，命名为openEuler。
- openEuler是一款开源操作系统。当前openEuler内核源于Linux，支持鲲鹏及其它多种处理器，能够充分释放计算芯片的潜能，是由全球开源贡献者构建的高效、稳定、安全的开源操作系统。openEuler适用于数据库、大数据、云计算、人工智能等应用场景。同时，openEuler也是一个面向全球的操作系统开源社区名称。



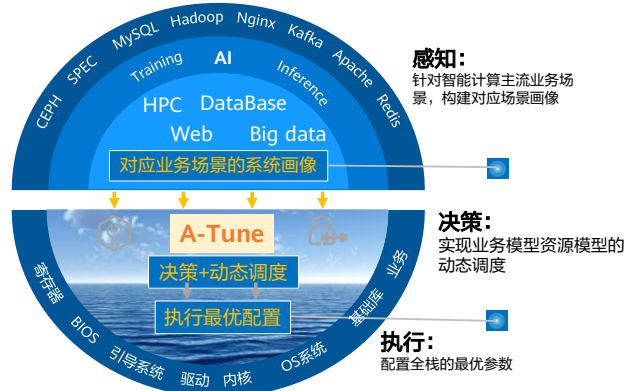
- 欧拉操作系统可广泛部署于服务器、云计算、边缘计算、嵌入式等各种形态设备，应用场景覆盖IT（Information Technology）、CT（Communication Technology）和OT（Operational Technology），实现统一操作系统支持多设备，应用一次开发覆盖全场景。

openEuler发展历程



openEuler技术特性：A-Tune资源调优自动化

- A-Tune是一种通过非侵入式系统画像的负载感知方法，识别业务并匹配最佳资源模型，实时响应业务特征变化的AI自动调优系统。



- 识别鲲鹏主流行业场景，A-Tune基于系统画像完成业务负载归类，实现资源模型建模和系统自优化，发挥鲲鹏算力。
- 基于机器学习完成系统画像，结合聚类+分类多次迭代实现业务负载的精确归类，涵盖当前ICT业务主流计算场景。
- 基于业务特征完成系统资源建模，合理优化系统资源及参数配置，实现系统能力的价值最大化。

openEuler技术特性：iSulad容器解决方案（1）

- iSula通用容器引擎（iSulad）是一种新的容器解决方案，具有以下特点：

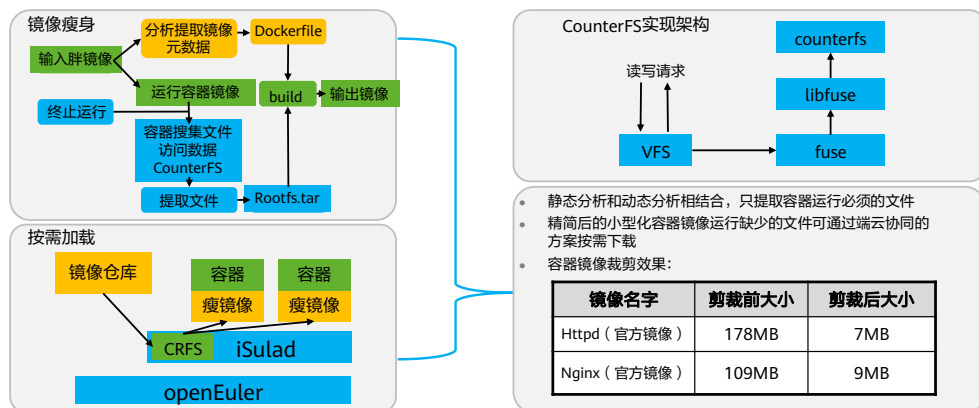
- 快速灵活&轻量级
- 可信&安全启动
- 升级不中断业务
- 增强安全性和调测特性



- iSula通用容器引擎（iSulad）是一种新的容器解决方案，提供统一的架构设计来满足CT和IT领域的不同需求。相比Golang编写的Docker，轻量级容器具有轻、灵、巧、快的特点，不受硬件规格和架构的限制，底噪开销更小，可应用领域更为广泛。openEuler-20.03-LTS中提供容器运行的基础平台iSula。
- 容器根据部署模型可以分成三种模型：应用容器、安全容器、系统容器。
 - 应用容器：应用最广泛的容器形态，容器中仅运行客户定义的应用程序。openEuler-20.03-LTS 集成moby 18.09版本，并在版本基础上进行了bugfix和稳定性增强。
 - 安全容器：本质上是虚拟机，但是接口封装为容器接口。openEuler-20.03-LTS 集成kata-container1.7版本，在社区版本的基础上，进行了稳定性和可靠性加固，增加了异构计算支持，更适于生产环境的Host CGroup配置与 DPDK/SPDK 高性能网络、存储支持。
 - 系统容器：本质上是容器，但是使用方式和虚拟机相同；着重解决虚拟机的厚重问题。iSula容器平台支持创建系统容器，并能支持在系统容器内动态调整设备、运行资源，且提供更优秀的user namespace隔离。

openEuler技术特性：iSulad容器解决方案（2）

- iSulad容器基础镜像支持按需剪裁，实现极致小型化。



openEuler其他技术特性概览

系统安装

openEuler-22.03-LTS可以让用户使用kickstart工具进行系统的自动化安装。

场景融合

openEuler-22.03-LTS是欧拉首个支持全景融合的社区长周期版本，满足服务器、云计算、边缘计算和嵌入式四大场景的多种不同类型设备部署要求和应用场景。

创新特性

openEuler 22.03-LTS合入了openEuler三个创新版中经过商业验证的创新特性，并针对四大场景首次发布新增特性，共新增代码2300 万行。

系统管理

openEuler-22.03-LTS使用systemd进行系统和服务的管理，systemd与SysV和Linux标准的init脚本兼容。



系统安全

openEuler-22.03-LTS提供多重安全手段，包括身份识别与认证、安全协议、强制访问控制、完整性保护、安全审计等安全机制，保障操作系统的安全性，为各类上层应用提供安全基础。

系统分析与调优

openEuler-22.03-LTS支持包含MySQL性能调优和大数据调优，包括如spark调优、hive调优、hbase调优。

编译器

openEuler-22.03-LTS支持使用毕昇编译器。毕昇编译器是华为编译器实验室针对鲲鹏等通用处理器架构场景，打造的一款高性能、高可信及易扩展的编译器工具链，增强和引入了多种编译优化技术，支持C/C++/Fortran等编程语言。

内核

openEuler-22.03-LTS采用kernel版本5.10。Linux Kernel 5.10是一个LTS 版本，提供长期支持并持续更新版本，以保障用户的Linux操作系统安全、可靠。

openEuler开源社区



01

开发者

- openEuler社区采用gitee作为开发系统。
- 开发者可以在开源社区提交bug、参与讨论、贡献代码文案, 也可以为openEuler的技术发展提出宝贵的建议, 共同建设openEuler的繁荣生态。

02

文档

- <https://openeuler.org>提供openEuler的所有文档, 包括发行说明、安装指南、管理员指南, 以及openEuler关键技术描述等文档。任何人均可以在<https://openeuler.org>上查看相关文档, 学习openEuler。

03

下载

- <https://openeuler.org>提供openEuler相关操作系统的下载, 有需求的玩家可以在开源社区下载openEuler相关操作系统。

04

社区会员

- openEuler开源社区欢迎各公司或者组织加入社区, 共同维护和管理社区, 共同打造操作系统生态圈。

- openEuler开源社区网址是: <https://www.openeuler.org/zh/>。

openGauss

- openGauss是一款携手伙伴共同打造全球领先的企业级开源关系型数据库，提供面向多核的极致性能、全链路的业务和数据安全、基于AI的调优和高效运维的能力。openGauss全面友好开放，采用木兰宽松许可证v2发行，深度融合华为在数据库领域多年的研发经验，结合企业级场景需求，持续构建竞争力特性。



- 多种存储模式支持复合业务场景。
- NUMA化数据结构支持高性能。
- 主备模式，CRC校验支持高可用。

openGauss发展历程



2001 - 2011	2011 - 2019	2019 - 2020	2021 - future
企业级内存数据库	- G行核心数据仓库、DWS华为云商用 - Z行核心业务系统替换商业数据库 - 支撑公司内部40+主力产品，在全球70+运营商规模商用3万+套，服务全球20+亿人口	2019.5.15 GaussDB全球发布 - 构筑合作伙伴生态 - 兼容行业主流生态，完成金融等行业对接 2020.6.30 openGauss开源	- 分享企业级数据管理能力 - 引领生态建设 - 促进数据库教育事业发展

高性能

- 提供了面向多核架构的并发控制技术结合鲲鹏硬件优化，在两路鲲鹏下TPCC Benchmark达成性能150万tpmc。
- 针对当前硬件多核的架构趋势，在内核关键结构上采用了NUMA-Aware的数据结构。
- 提供Sql-bypass智能快速引擎技术。

其他特性



高可用

- 支持主备同步、异步以及级联备机多种部署模式。
- 数据页CRC校验，损坏数据页通过备机自动修复。
- 备机并行恢复，10秒内可升主提供服务。



高安全

- 支持全密态计算、访问控制、加密认证、数据库审计、动态数据脱敏等安全特性，提供全方位端到端的数据安全保护。



易运维

- 基于AI的智能参数调优和索引推荐，提供AI自动参数推荐。
- 慢SQL诊断，多维性能自监控视图，实施掌控系统的性能表现。
- 提供在线自学习的SQL时间预测。



全开发

- 采用木兰宽松许可证协议，允许对代码自由修改、使用、引用。
- 数据库内核能力全开放。
- 提供丰富的伙伴认证、培训体系和高校课程。

目录

1. 鲲鹏计算产业概述

- 什么是鲲鹏
- 鲲鹏处理器
- 鲲鹏计算整机
- 鲲鹏生态
- 鲲鹏HPC解决方案

2. 昇腾计算产业概述

什么是HPC

- **HPC: High Performance Computing, 高性能计算**, 指使用很多处理器（作为单个机器的一部分）或者某一集群中组织的几台计算机（作为单个计算资源操作）的计算系统 and 环境，通常是将大规模运算任务拆分并分发到各个服务器上进行并行运算，再将计算结果汇总得到最终结果。



- HPC发展已经有半个多世纪，早期HPC基于定制计算机，专用于基础科学研究和国防战略等特定领域，随着技术的发展，现在绝大部分HPC基于x86服务器的集群架构进行构建。
- HPC计算集群主要由计算服务器、共享存储、高速互连设备以及配套的管理软件、平台软件和应用软件构成。

理解高性能计算



气候和天气预报模拟



教育科学



制造业及材料科学



生命科学，药品开发

HPC有助于减少新车研发时间

60个月->24个月



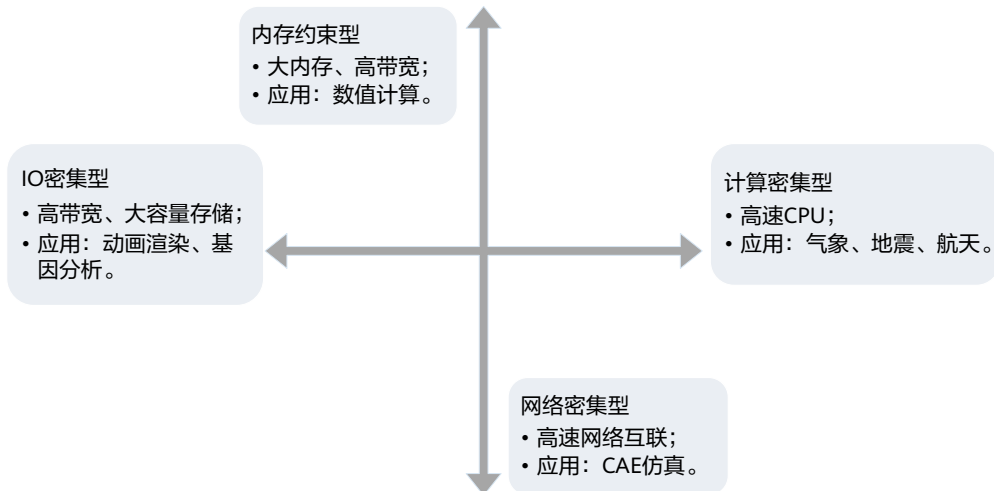
HPC有助于加快新药研发鉴定

普通PC计算5000年
->高性能计算100天



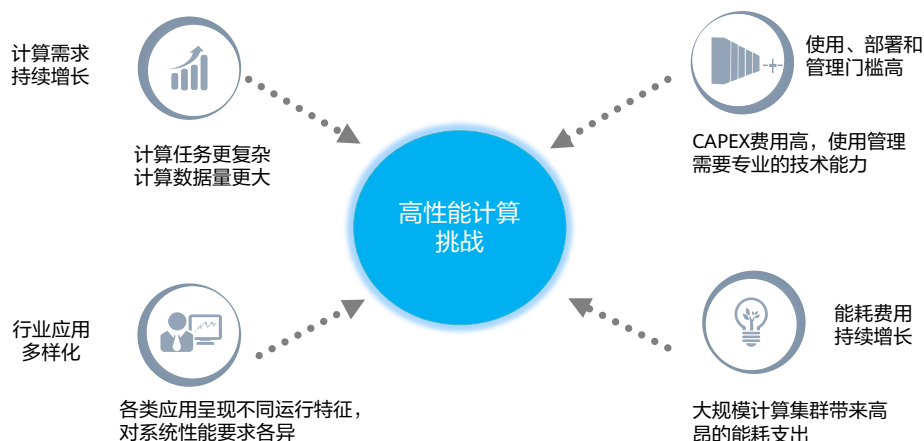
- 高性能计算，是通过高速网络将大量服务器进行互联形成计算机集群，与高性能存储一起，求解科研、工业界最复杂的科学计算问题（科学研究领域三大范式：理论科学，实验科学，计算科学）。

HPC主要应用场景特点



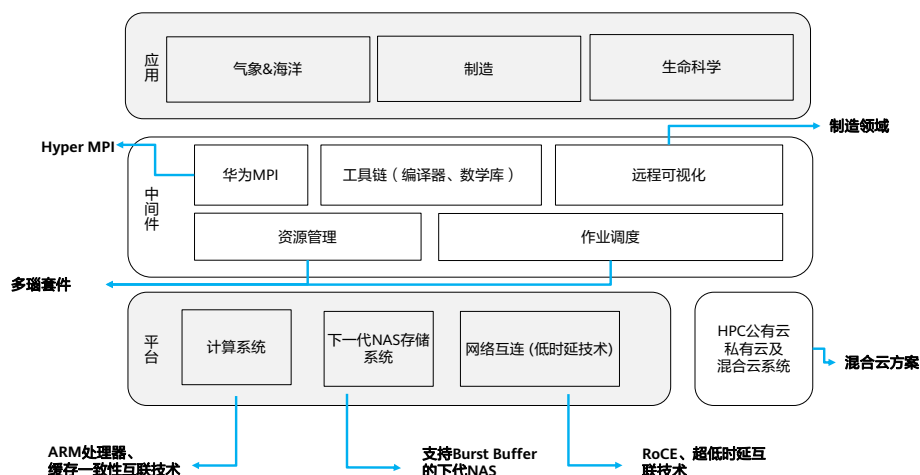
- 不同的行业和研究领域在应用高性能计算时，对高性能计算集群有不同的系统要求：
 - 譬如在涉及天体物理或分子动力模拟计算应用场景时，需要处理大模型计算，多处理器、共享大容量内存的配置允许中间数据存储在内存中，CPU可以直接从内存中获取而不需要和硬盘交换数据，可以减少IO频度，明显提高计算速度。
 - 在气象预报应用场景中，计算本身有时效性要求，应用的计算量大、并行要求高，所以强劲的计算能力有利于提升整体的计算效率。
 - 在汽车、航空、芯片制造领域借助高性能计算进行产品的CAE仿真设计，需要对复杂的产品几何模型进行求解，各节点在计算过程中有频繁的通讯需求。为了避免CPU等待数据进行计算的情况，因此需要低时延、高带宽的网络来连接大量的计算节点间的数据传输，提升计算效率。
 - 在基因分析场景，序列比对、拼接需要对海量、复杂、多变的数据进行分析计算，单次测序数据量的大幅度提升，在计算过程中会产生大量的数据，因此需要配置高带宽、大容量的共享存储，满足计算节点频繁的数据IO读写请求。

高性能计算当前挑战



- 当前高性能计算在各行业应用中面对以下几个挑战：
 - 计算机辅助仿真和模拟应用越来越广泛，对高性能计算的资源需求越来越大，需要考虑采用什么样的系统架构和技术，实现更大规模的集群。
 - 各行业有各自的HPC应用软件，对集群的性能和配置需求也各不相同，需要理解这些应用特征，为应用部署实现最佳的软件运行环境。
 - HPC集群的建设和使用门槛相对还是比较高的，数量众多的专有硬件成本非常高，建设更新周期也比较长，如何以更灵活的方式在各行业的应用普及HPC，是个具有挑战的课题。
 - 更大规模的集群，必然带来更高大的能耗，为了降低整体能耗和数据中心TCO，客户对集群的能耗和散热解决方案非常敏感。

华为HPC解决方案能力



- 基础软件：HPCKit提供通信库（Hyper MPI）、鲲鹏KML数学库、鲲鹏毕昇/GCC编译器组件的一键部署和最优协同。
 - Hyper MPI是基于Open MPI 4.1.1和Open UCX 1.10.1，支持MPI-V3.1标准的并行计算API接口，新增了优化的集合通信框架。同时，Hyper MPI对数据密集型和高性能计算提供了网络加速能力，使能了节点间高速通信网络和节点内共享内存机制，以及优化的集合通信算法。
 - 鲲鹏数学库（KML）是基于华为鲲鹏处理器深度优化的高性能数学计算加速库，支持快速傅里叶变化、矩阵计算、向量化运算、三角函数、对数等常用数学函数。
 - 毕昇编译器基于开源LLVM开发，并进行了优化和改进，同时支持Fortran语言前端，是针对鲲鹏平台的高性能编译器。除LLVM通用功能和优化外，对中端及后端的关键技术点进行了深度优化，并集成AutoTuner特性支持编译器自动调优。
- 多瑙套件：
 - 多瑙Portal：多瑙Portal是华为全自研的HPC集群管理平台，通过可视化界面为用户提供提供了便捷的HPC集群系统数据管理和软硬件资源管理功能，串联整个工作流程，帮助用户合理地进行作业调度和资源分配，提升集群系统计算能力利用率。
 - 多瑙调度器：多瑙调度器是华为全自研的HPC集群调度器，提供大规模集群下的高资源利用率、高吞吐量的作业调度能力。

HPC解决方案性价比高

- 基因测序、流体力学、气象环保主要为内存敏感型应用以及整型计算（或者单精度计算）为主，适合鲲鹏云HPC应用场景。



- 性价比领先：
 - 独家发布基于自研鲲鹏芯片的HPC解决方案，性价比优于传统方案30%。
 - 并行文件系统PFS，同时支持对象存储和POSIX接口，替代传统并行文件系统，数据免拷贝，TCO降低20%。
 - 独家支持100G InfiniBand高速网络，时延低至1.5微秒，支持RDMA，性能业界领先。
- 易部署、易使用、易维护：
 - 提供云上HPC集群管理服务，支持集群一键式发放、快速部署，效率提升90%。
 - 提供集群管理、作业管理、节点管理等可视化HPC集群管理能力。
 - 持续集成HPC编译器、MPI等各种开源商用HPC中间件及业务软件。
- 丰富的HPC生态：
 - 基因测序：GATK、bwa、Bowtie、Blast等近百款软件已成功移植。
 - 气象环保：WRF、ROMS、CAMX、CMAQ等近十款软件已成功移植。
 - 流体力学：OpenFORM、SU2等已经成功移植。

某大学的“超级计算系统”



业务挑战

- 科研人员作业排队时间太长，用户体验差，影响科研进度；
- 原有高性能计算平台利用率过高，无法进行大规模并行计算，制约科研广度、深度、精度；
- 存储小、性能低；
- 受现有机房空间及电力限制，普通的风冷服务器无法放下，希望选择液冷服务器提高能效实现节能。

解决方案

- 该“超级计算系统”采用华为全液冷的解决方案和高密服务器，规模高达30480颗CPU核，系统理论峰值性能2.52 Pflops，实测HPL计算效率为63.95%，甚至部分计算节点HPL效率能跑出71%以上的超高性能。
- 该“超级计算系统”也为科研用户提供了多种算力平台，采用了20台华为TaiShan服务器构建基于鲲鹏架构的计算集群，为生物信息学计算软件和流体力学提供超强算力。

- 华为助力某大学打造千万亿次“HH超级计算系统”。
- 客户背景：
 - 客户为一所以前沿科学和高新技术为主的综合性全国重点大学。在量子信息、化学与材料、物理、微尺度物质科学、核科学等学科领域，位居国内科研前沿，此类学科对于计算量需求巨大，也就形成了高性能计算天然的应用场景。
- 客户价值：
 - “HH超级计算系统”已成为当前中国高校中最大高性能计算校级平台之一，稳定高效运行近一年，完成几十万个作业，超2亿CPU核小时，支持了大规模的天体宇宙模拟、高并行的量化计算与等离子体模拟等帮助该大学完成了一系列学术科研突破；
 - “HH超级计算系统”采用720台华为液冷高密服务器，仅需10个机柜即可部署，不仅大幅节省部署空间，同时实现绿色节能，相比传统服务器每年至少可节省电费17余万元。

目录

1. 鲲鹏计算产业概述

2. 昇腾计算产业概述

■ 什么是昇腾

- 昇腾硬件
- 昇腾软件
- 昇腾行业应用
- 华为人工智能计算中心

什么是昇腾



- 昇腾计算产业是基于昇腾系列（HUAWEI Ascend）处理器和基础软件构建的全栈 AI 计算基础设施、行业应用及服务，包括昇腾系列处理器、系列硬件、CANN（Compute Architecture for Neural Networks，异构计算架构）、AI计算框架、应用使能、开发工具链、管理运维工具、行业应用及服务全产业链。

目录

1. 鲲鹏计算产业概述

2. 昇腾计算产业概述

- 什么是昇腾
 - 昇腾硬件
- 昇腾软件
- 昇腾行业应用
- 华为人工智能计算中心

昇腾AI处理器 – 310和910



- Ascend-Mini
- 架构: 达芬奇
- 最大功耗: 8 W

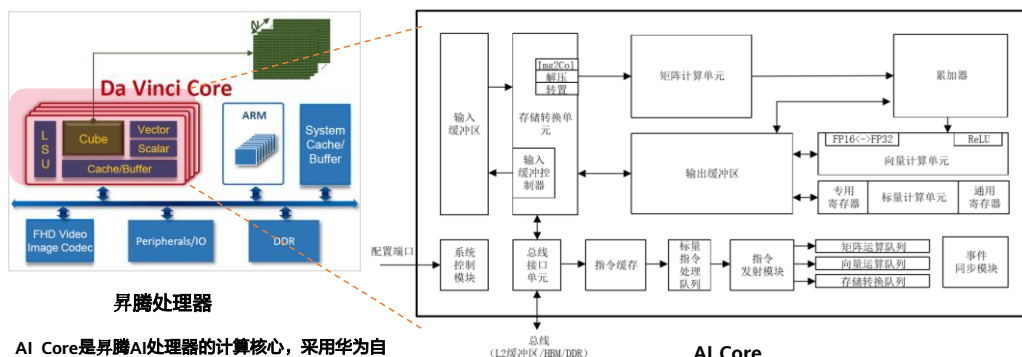
面向推理场景



- Ascend-Max
- 架构: 达芬奇
- 最大功耗: 310 W

面向训练场景

AI Core：在昇腾处理器中的位置



- AI Core是昇腾AI处理器的计算核心，采用华为自主研发的达芬奇架构，通常也被叫做Da Vinci Core。
- 根据不同的处理器版本，AI Core里的计算、存储和带宽资源有不同的规格，在本课程中，主要是按照Ascend310/910来介绍。

达芬奇架构主要部分：

- 计算单元：包含三种基础计算资源（矩阵计算单元、向量计算单元、标量计算单元）。
- 存储系统：AI Core的片上存储单元和相应的数据通路构成了存储系统。
- 控制单元：整个计算过程提供了指令控制，相当于AI Core的司令部，负责整个AI Core的运行。

- 计算单元：三种基础计算资源：矩阵计算单元、向量计算单元和标量计算单元，分别对应矩阵、向量和标量三种常见的计算模式。
 - 矩阵计算单元（Cube Unit）：矩阵计算单元和累加器主要完成矩阵相关运算。一拍完成一个fp16的 16x16与16x16矩阵乘（4096）；如果是int8输入，则一拍完成 16*32 与 32*16 矩阵乘（8192）。
 - 向量计算单元（Vector Unit）：实现向量和标量，或双向量之间的计算，功能覆盖各种基本的计算类型和许多定制的计算类型，主要包括FP16/FP32/Int32/Int8等数据类型的计算。一拍可以完成两个128长度fp16类型的向量相加/乘，或者64个fp32/int32类型的向量相加/乘。
 - 标量计算单元（Scalar Unit）：相当于一个微型CPU，控制整个AI Core的运行，完成整个程序的循环控制、分支判断，可以为Cube/Vector提供数据地址和相关参数的计算，以及基本的算术运算。
- 存储系统：AI Core采用了大容量的片上缓冲区设计，通过增大的片上缓存数据量来减少数据从片外存储系统搬运到AI Core中的频次，从而可以降低数据搬运过程中所产生的功耗，有效控制了整体计算的能耗。
 - 输入缓冲区：用来暂时保留需要频繁重复使用的数据，不需要每次都通过总线接口到AI Core的外部读取，从而在减少总线上数据访问频次的同时也降低了总线上产生拥堵的风险，达到节省功耗、提高性能的效果。
 - 输出缓冲区：用来存放神经网络中每层计算的中间结果，从而在进入下一层计算时方便的获取数据。相比较通过总线读取数据的带宽低，延迟大，通过输出缓冲区可以大大提升计算效率。
- 控制单元：
 - 控制单元主要组成部分为系统控制模块、指令缓存、标量指令处理队列、指令

发射模块、矩阵运算队列、向量运算队列、存储转换队列和事件同步模块。

Atlas人工智能计算平台全景图



- 华为Atlas人工智能计算平台基于华为昇腾系列AI处理器，通过模块、板卡、小站、服务器、集群等丰富的产品形态，打造面向“端、边、云”的全场景AI基础设施方案。作为华为全栈AI解决方案的重要组成部分，Atlas在推出推理产品的基础上，通过训练产品，为业界呈现一个完整的AI解决方案。同时，我们还通过全场景部署，打通云边端协同，让AI能力赋能各环节。

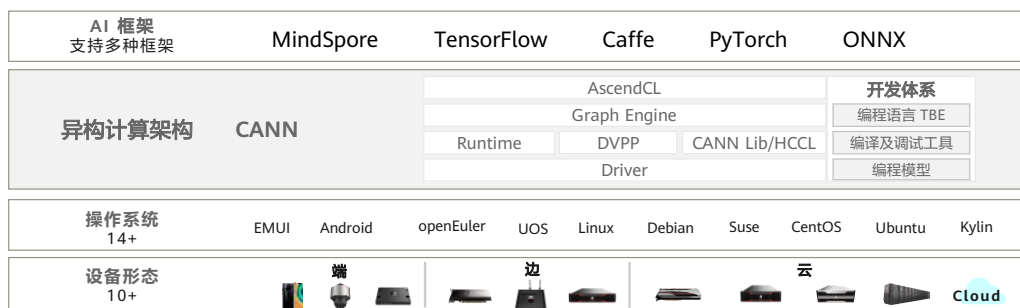
目录

1. 鲲鹏计算产业概述

2. 昇腾计算产业概述

- 什么是昇腾
- 昇腾硬件
- 昇腾软件
- 昇腾行业应用
- 华为人工智能计算中心

CANN：端边云协同，使能全场景AI开发



端边云全场景协同

10+ 设备形态
14+ 操作系统
多种AI框架



AscendCL使能高效开发

统一API，后向兼容
四大开放性设计
两种算子开发方式

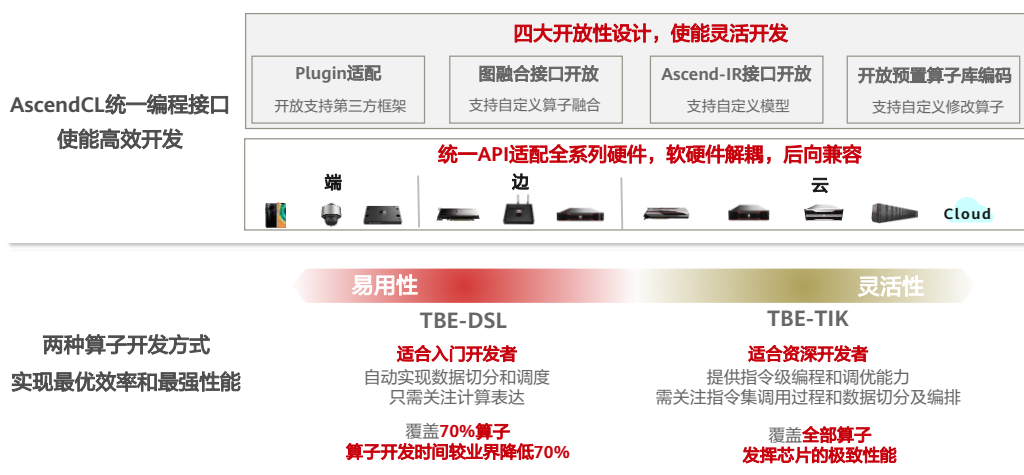


释放硬件澎湃算力

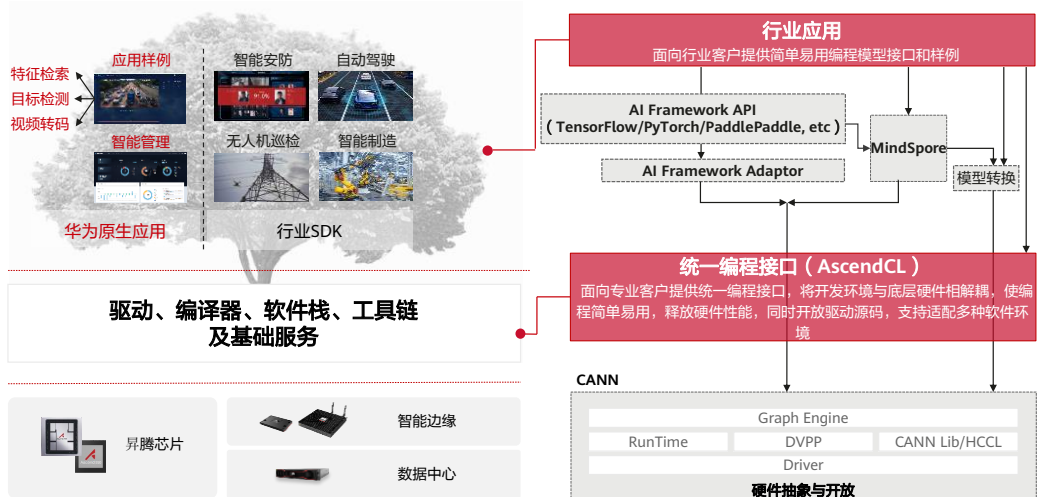
极致的图编译技术
丰富的算子库支持，1000+高性能算子

- 应用开发框架AscendCL：提供Device管理、Context管理、Stream管理、内存管理、模型加载与执行、算子加载与执行、媒体数据处理等C++ API库供用户开发深度神经网络应用，通过加载模型推理实现目标识别、图像分类等功能。
- 计算图引擎GE：GE作为图编译和运行的控制中心，提供运行环境管理、执行引擎管理、算子库管理、子图优化管理、图操作管理和图执行控制。
- 数字视觉预处理DVPP：主要实现视频解码（VDEC）、视频编码（VENC）、JPEG编解码（JPEGD/E）、VPC（预处理）。
- 集合通信算子库HCCL：集合通信对外提供集合通信算子，支持网卡及集群不同节点间的RoCE传输功能，为分布式训练中不同npu之间提供高效的数据传输能力。

CANN：灵活高效，使能全场景AI开发



CANN：能力分层开放，使能行业AI应用



MindSpore：自研全场景AI计算框架



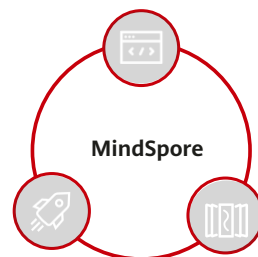
- 基于上面的三个设计理念，我们研发了自己的MindSpore的核心架构，核心分为四层：按需协作分布式架构、调度、分布式部署、通信库都在同一层，再往下是执行高效层（包括数据模型下沉），另外是并行的Pipeline执行层和深度图的优化、算子的融合，再上层有MindSpore原生的计算图表达。支持了自动微分、自动并行、自动调优的特性点，支撑了符合我们的设计理念全场景API：开发算法即代码、运行高效、部署态灵活形成一体化。

MindSpore：三大创新助力全场景AI应用



全流程极简

从小白到专家，打造简单易用的AI开发套件



全自行并行

模型、算子自动切分，多核自动扩展到集群

全场景协同

自适应部署多样化设备，模型“私人订制”

MindX: “2+1+X”

2 个组件

1 模型库

X 行业应用



深度学习平台
基础组件



智能边缘平台
基础组件



预训练高性能模型



行业AI应用软件开发套件

MindX DL深度学习组件：开放生态、开源参考架构

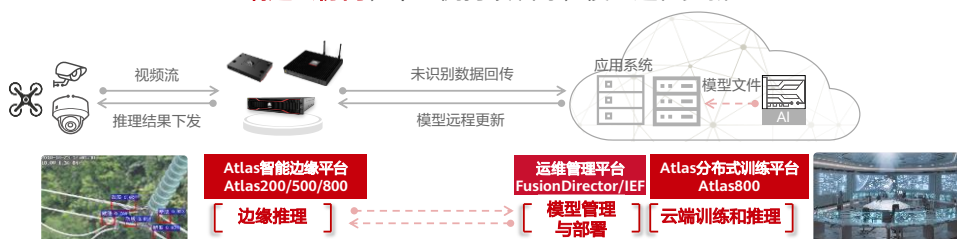


- MindX DL是kubernetes生态下，支持Atlas 800训练服务器、Atlas 800推理服务器等包含华为AI处理器的服务器（标卡和模组形态）必备组件。
- 提供昇腾AI处理器资源注册和监控、昇腾AI处理器优化调度、分布式训练集合通信配置生成等基础功能，同时还兼容GPU服务器调度。
- 能够快速使能合作伙伴进行深度学习平台开发，降低合作伙伴使用门槛。

- Ascend Device Plugin：基于Kubernetes设备插件机制，增加昇腾AI处理器的设备发现、设备分配、设备健康状态上报功能，使Kubernetes可以管理昇腾AI处理器资源。
- HCCL-Controller：华为研发的一款用于昇腾AI处理器训练任务的组件，利用Kubernetes（简称K8s）的informer机制，持续记录训练任务及其Pod的各种事件，并读取Pod的昇腾AI处理器信息，生成对应的configmap。该configmap包含了训练任务所依赖的集合通讯配置，方便训练任务更好地协同和调度底层的昇腾AI处理器，无需用户手动配置。
- Volcano：基于开源Volcano调度的插件机制，增加昇腾AI处理器的亲和性调度、故障重调度等特性，最大化发挥昇腾AI处理器计算性能。
- NPU-Exporter：该组件为Prometheus生态组件，提供了昇腾AI处理器资源各种指标的实时监测，可实时获取昇腾AI处理器利用率、温度、电压、内存，以及昇腾AI处理器在容器中的分配状况等信息。
- NodeD：提供节点状态上报功能，如节点心跳上报；为Atlas A2训练系列产品提供节点硬件故障上报功能，如CPU、内存等故障信息上报。
- Elastic-Agent：在数据并行、混合并行场景下提供策略恢复功能。
- Resilience-Controller：韧性控制器，提供弹性训练的韧性控制。在训练任务使用的硬件故障时，可以剔除该硬件，继续训练。
- Ascend Docker Runtime：容器引擎插件，为所有AI作业提供NPU容器化支持，使用户进行AI作业时能够以Docker容器的方式平滑地在昇腾设备上运行。
- Ascend Operator：支持MindSpore、PyTorch、TensorFlow三个AI框架在Kubernetes上进行分布式训练的插件，CRD（Custom Resource Definition）中定义了AscendJob任务，用户只需配置yaml文件，即可以为不同AI框架的分布式训练任务提供相应环境变量。

云边协同，发挥全场景协同优势

端边云协同，中心侧持续训练，模型远程更新



统一开发

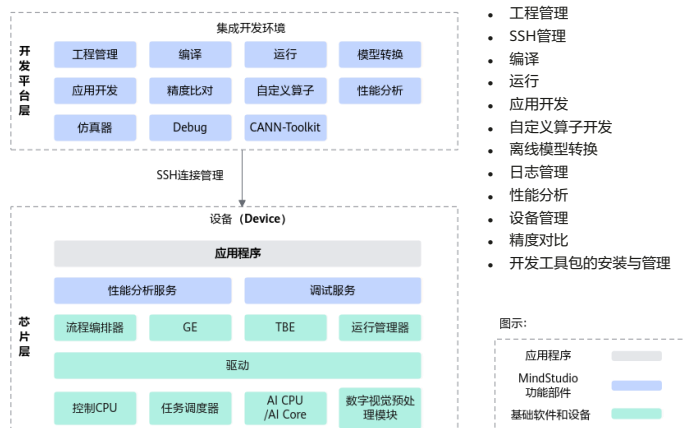
统一运维

安全升级

Atlas 业界方案	- 基于达芬奇和CANN的统一开发架构，一次开发，端边云皆可使用 - 边缘侧和中心侧分别使用不同开发架构，模型无法自由流动，需二次开发	- 管理最多5万个节点，实现中心边缘设备统一管理，模型推送/设备升级等均可远程完成 无运维管理工具，仅开放API，客户需自行开发	- 传输通道安全加密 模型加密，双重保障 无加解密引擎，对模型未加密
---------------------------------	--	--	--

MindStudio架构

- MindStudio是一套基于IntelliJ框架的开发工具链平台，提供了应用开发、调试、模型转换功能，同时还提供了网络移植、优化和分析功能，为用户开发应用程序带来了极大的便利。



- Mind Studio功能框架如图所示，目前含有的工具链包括：工程管理工作、编译工具、流程编排工具、离线模型工具、比对工具、日志管理工具、自定义算子工具、性能分析工具、设备管理工具、设备开发工具包等多种工具。
- 针对算子开发：**提供全套的算子开发、调优能力。通过Mind Studio提供的工具链也可以进行第三方算子开发，降低了算子开发的门槛，并提高算子开发及调试调优的效率，有效提升了产品竞争力。
- 针对网络模型的开发：**集成了离线模型转换工具、模型量化工具、模型精度比对工具、模型运行性能分析工具、日志分析工具，提升了网络模型移植、分析和优化的效率。
- 针对计算引擎开发：**预置了典型的分类网络、检测网络等计算引擎代码，降低了开发者的技术门槛，加快了开发者对AI算法引擎的编写及移植效率。
- 针对应用开发：**集成了各种工具如分析器（Profiler）和编译器（Compiler）等，为开发者提供了图形化的集成开发环境，通过Mind Studio能够进行工程管理、编译、调试、性能分析等全流程开发，能够很大程度提高开发效率。

目录

1. 鲲鹏计算产业概述

2. 昇腾计算产业概述

- 什么是昇腾
- 昇腾硬件
- 昇腾软件
- 昇腾行业应用
- 华为人工智能计算中心

行业SDK：凝聚行业经验，结合AI最佳实践，加速行业AI应用落地



电力：业界首创智能无人巡检，作业效率提升5倍



医疗：AI让医疗影像诊断效率提升50倍以上



制造：从人工质检到AI质检，从“中国制造”到“中国智造”

智能制造开发套件已在松山湖规模部署，开发部署效率提升10倍



< 智能制造开发套件 >

核心质检流程全面插件化，无需从头开发，直接进行流程编排



30k 行 代码开发量

60 人月 业务上线

200+ API接口调用

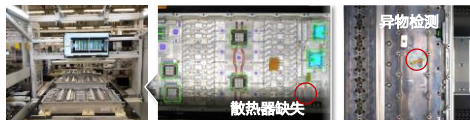


1.5k 行 代码开发量

3 人月 业务上线

6 API接口调用

松山湖产线AI工业质检 190+ 产线已应用，覆盖服务器、无线、终端等产品



质检准确率 99.9%

质检效率提升 3倍

交通：AI助力全国高速路网升级，通行效率提升5倍



目录

1. 鲲鹏计算产业概述

2. 昇腾计算产业概述

- 什么是昇腾
- 昇腾硬件
- 昇腾软件
- 昇腾行业应用
- 华为人工智能计算中心

以“一中心四平台”方案推动人工智能产业集约集聚发展



建设方式：政府统筹建设像“发电厂”一样的基础设施

政府统筹建设人工智能计算中心



普惠算力

承担新基建公共基础设施职责，
解决中小企业算力昂贵的问题



AI计算中心

集群计算

普惠公共算力



企业小型算力



板卡
单服务器

企业购买算力昂贵
小算力训练速度慢



安全可靠

依托政府公信力，构建安全可信
算力资源，并实现数据安全共享



共享数据集



AI计算中心

ISV

ISV

ISV

企业自建

ISV

企业自建

ISV

企业自建

ISV

企业自建

ISV



产业协同

充分发挥政府的产业调控作用，
实现人工智能与产业深度融合

政府调控

政

产

学

研

用

相互渗透，互为支撑，互为动力

产

研

产

用

技术与产业较割裂

产

用

产

学

研



节约能耗

集约化统一建设，节约能耗，
扎实做好碳达峰、碳中和工作



AI计算中心

集群计算

液冷 PUE < 1.1



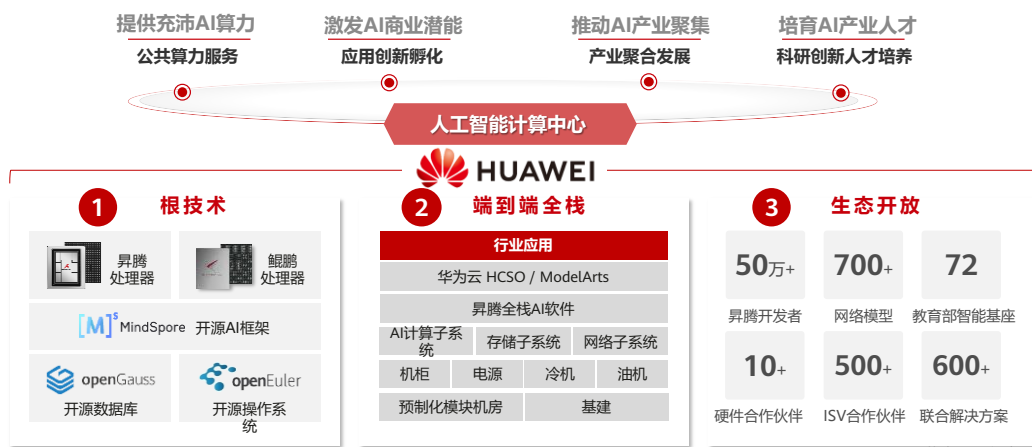
企业小型算力



服务器

风冷 一般PUE > 1.4

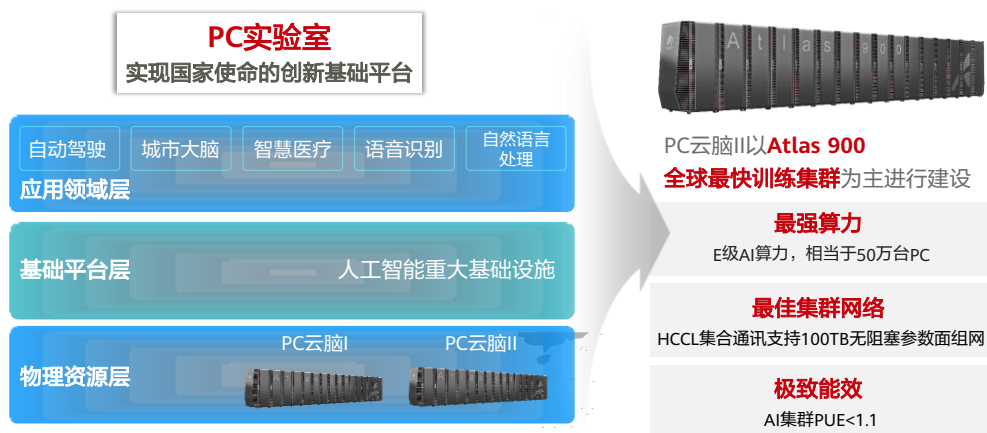
华为方案：自主可控、端到端全栈、生态开放的人工智能计算中心解决方案



截止2022年底



人工智能计算中心：Atlas助力打造国家级人工智能平台



思考题

1. （多选题）以下关于华为鲲鹏920处理器的描述，正确的是哪些项？（ ）
 - A. 采用了7nm的制造工艺
 - B. 支持8通道的DDR4控制器
 - C. 支持PCIe 4.0接口，并兼容PCIe 3.0/2.0/1.0
 - D. 支持多种加速器
2. （多选题）华为AI全栈解决方案包括？（ ）
 - A. 芯片层
 - B. CANN层
 - C. 框架层
 - D. 应用使能层

- 答案：
- ABCD
- ABCD

本章总结

- 本章我们深入学习了什么是鲲鹏，熟悉了鲲鹏芯片和整机产品同时也掌握鲲鹏生态体系内容，比如openEuler和openGauss的特点及优势。学习了鲲鹏HPC解决方案。
- 同时还深入学习了什么是昇腾，熟悉了昇腾硬件架构及其硬件产品，还学习了昇腾AI计算平台CANN和其他昇腾软件产品。了解了昇腾行业应用和华为人工智能计算中心。

缩略语

- CPU: Central Processing Unit , 中央处理器, 是计算机的主要设备之一, 其功能是解释计算机指令以及处理计算机软件中的数据, 与内部存储器、输入及输出设备成为现代电脑的三大部件神经处理单元。
- GPU: Graphics Processing Unit, 图形处理器, 是一种专门在个人电脑、工作站、游戏机等设备上图像运算工作的微处理器, 可用于桌面的图片渲染。
- CISC: Complex Instruction Set Computer, 复杂指令集计算机, 是一种微处理器指令集架构, 每个指令可执行若干低端操作, 诸如从存储器读取、存储、和计算操作, 全部集于单一指令之中。
- RISC: Reduced Instruction Set Computer, 精简指令集计算机, 是计算机中央处理器的一种设计模式。
- NVMe : Non-Volatile Memory Express, 非易失性存储器标准, 是一种用于高度并行数据传输的协议, 可减少闪存存储和固态硬盘 (SSD) 中使用的每个输入/输出 (I/O) 的系统开销。
- RAID : Redundant Array of Independent Disks, 独立硬盘冗余阵列;

缩略语

- ISV: Independent Software Vendor, 独立软件供应商, 这类公司通常为特定行业领域提供软件支持;
- C&SI: Consulting and Systems Integration, 咨询与系统集成;
- IHV: Independent Hardware Vendor, 独立硬件商, 这类公司通常为特定行业领域提供软件支持;
- SDK: Software Development Kit, 软件开发套件, 是通常由硬件平台、操作系统 (OS) 或编程语言的制造商提供的一套工具。

Thank you.

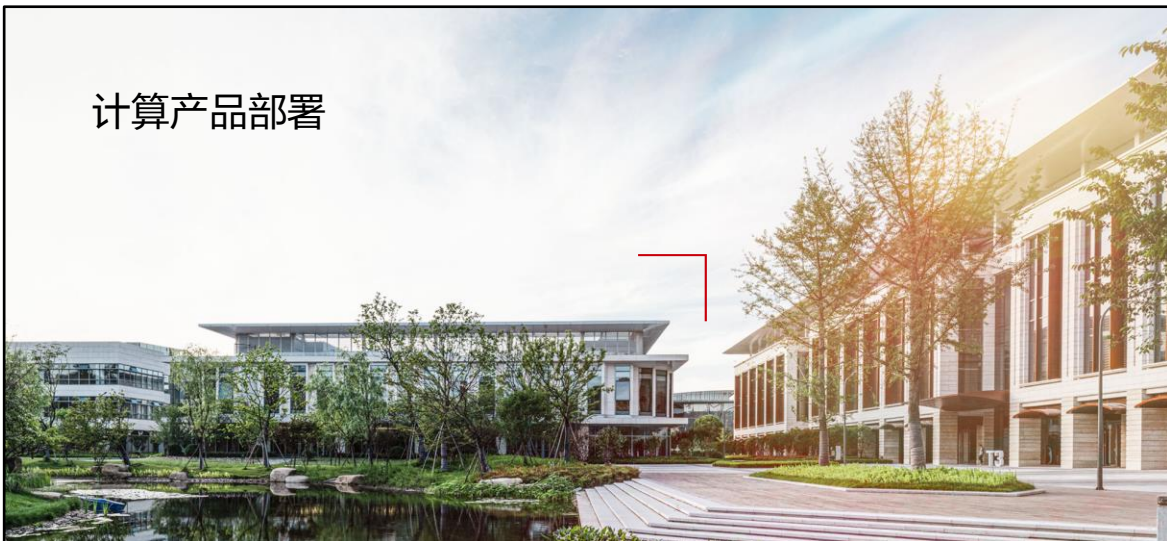
把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。
Bring digital to every person, home, and
organization for a fully connected,
intelligent world.

Copyright©2024 Huawei Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.



计算产品部署



前言

- 本章主要讲述服务器安装部署的环境条件及注意事项，并介绍了以Atlas800-3000为例，进行安装部署的全部流程。

目标

- 学完本课程后，您将能够：
 - 了解常见的机房结构
 - 熟悉服务器安装部署要求
 - 熟悉服务器部署的关键步骤及配置方式
 - 了解如何在服务器上部署业务

目录

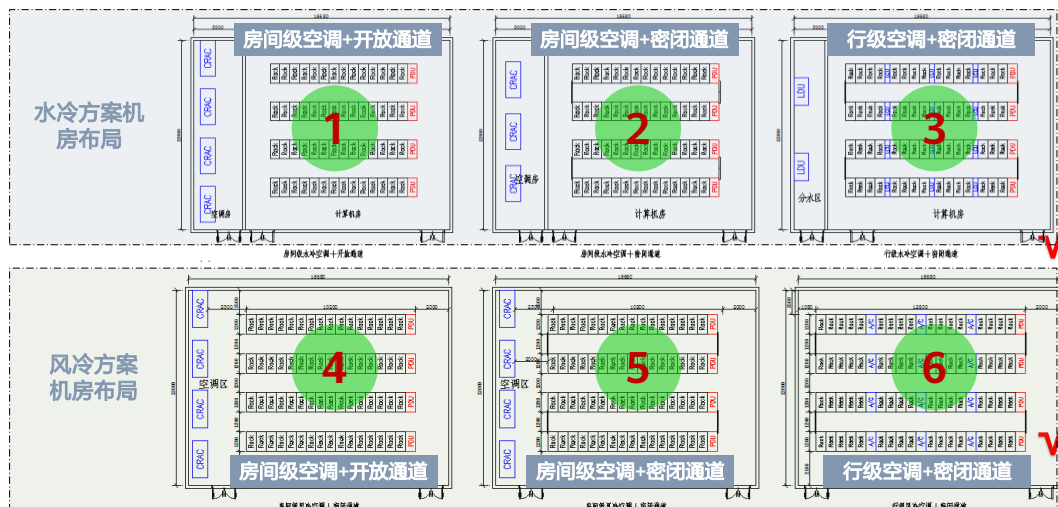
1. 常见的机房结构
2. 安装部署要求
3. 部署关键步骤及配置项
4. 业务部署

机房存在的必要性

- 企业想要部署业务系统，只有服务器是不够的，还需要一系列的基础设施和运维管理环境。
- 机房是信息时代的重要基础设施，承担着保障企业数据安全、提高运营效率等重要任务：
 - 提供高速、稳定的网络环境，便于企业进行远程协作和启用服务；
 - 提供可靠的存储和备份设施，确保企业数据的安全性和完整性；
 - 通过集中管理和维护，降低企业的IT运营成本和风险；
 - 通过智能监控和自动化管理，提高企业的生产效率和响应速度。



常见的机房布局

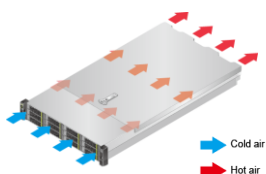


- 行级空调+密闭通道方案无论在**机房利用率**、支持最大单机柜**功率密度**、**PUE**、**安全性**方面均为三种方案中最优；房间级+密闭通道方案次之，最差的是房间级空调+开放通道建设模式。

目录

1. 常见的机房结构
- 2. 安装部署要求**
3. 部署关键步骤及配置项
4. 业务部署

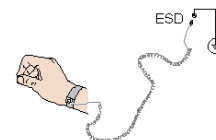
Atlas 800的安装环境要求



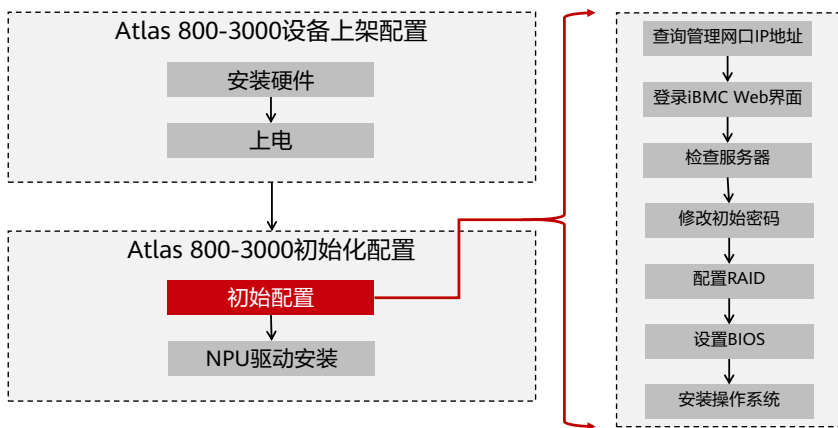
- 机柜要求：
 - 满足IEC297标准的宽19英寸、深1000 mm以上的通用机柜。
 - 在机柜门上安装防尘网；在机柜后面提供交流电源接入。
- 温度与湿度要求：
 - 温度：5°C ~ 40°C（41°F ~ 104°F）。
 - 湿度：8% RH ~ 90% RH（无冷凝）。
- 空间与通风要求：
 - 服务器必须安装在出入受限区域；保持设备所在区域整洁；确保机柜前后都要空余800 mm的空间。
 - 服务器入风口处应避免有障碍物阻挡，影响正常进风和散热；服务器放置位置的空调送风量应足够提供服务器需要的风量，保证服务器内部各器件散热。
 - 服务器从前面板吸入凉风，从后面板排出热风，具体的散热气流走向如图，前后方都必须通风良好。

部署规范介绍 – 防静电

- 静电可能对产品造成损伤的，请注意以下操作准则：
 - 所有机房应该铺设防静电地板，使用防静电工作椅。机房的隔板、屏风、窗帘等应使用防静电材料。
 - 机房的落地式用电设备、金属框架、机架的金属外壳必须直接与大地连接，工作台上的所有用电仪器工具应通过工作台的公共接地点接地。
 - 请注意监控机房温度、湿度。暖气会降低室内湿度并增加静电。
 - 在运输、保管服务器组件的过程中，必须使用专用的防静电袋与防静电盒，以确保服务器组件的防静电安全。
 - 机房内的人员在进行服务器组件安装、插拔等接触操作时必须佩戴防静电腕带，并将接地端插入机架上的ESD插孔。
 - 在接触设备前，应当穿上防静电工作服、佩戴防静电手套或防静电腕带、去除身体上携带的易导电物体（如首饰、手表等），以免被电击或灼伤。
 - 防静电腕带的两端必须接触良好，一端接触您的皮肤，另一端牢固地连接到机箱的ESD接口。
 - 在更换的过程中，应将所有还没有安装的服务器组件保留在带有防静电屏蔽功能的包装袋中。
将暂时拆下来的服务器组件放置在具有防静电功能的泡沫塑料垫上。
 - 请勿触摸焊接点、引脚或裸露的电路。



安装步骤概览



安装服务器

- 搬运时双手托住服务器底端靠中部位置，从两侧抬起服务器，并将服务器放置在滑道上。
- 将服务器推入机柜，如下图中①所示。
- 将服务器两侧挂耳紧贴机柜方孔条，拧紧挂耳上的松不脱螺钉，固定服务器，如下图中②所示。



上电

- 根据实际需求连接VGA线缆、USB线缆、灵活网卡网线、管理网口线缆和串口线缆。
- 链接电源模块线缆。



- 打开机柜供电开关。服务器连接的外部电源空气开关电流规格：交流电源32 A或直流电源63 A。
- 将电源模块上电，根据系统默认策略，服务器会随电源模块一起上电。
 - 系统默认“通电开机策略”为“保持上电”，即服务器的电源模块通电后系统自动开机，用户可在 iBMC WebUI 的“电源与能耗 > 电源控制”界面进行修改。
 - 若电源模块已上电，服务器不自动开机，可以依据产品文档自行上电。

目录

1. 常见的机房结构
2. 安装部署要求
- 3. 部署关键步骤及配置项**
 - 服务器初始化配置
 - BIOS关键参数及配置
 - RAID配置
 - 操作系统安装
 - NPU驱动安装
4. 业务部署

Atlas 800 – 3000的默认数据

类别	名称	默认值	备注
iBMC管理系统网口数据	管理网口IP地址与子网掩码	默认IP地址: 192.168.2.100 默认子网掩码: 255.255.255.0	进入BIOS界面后可查看IP地址
iBMC管理系统默认登录用户	用户名与密码	用户名: root 默认密码: Huawei12#\$	iBMC所有用户名及密码以《Atlas硬件产品 用户清单》为准
iBMC Web界面或iBMC命令行登录	用户名与密码	用户名: Administrator 默认密码: Admin@9000	
iBMC U-boot数据	默认密码	默认密码: Admin@9000	

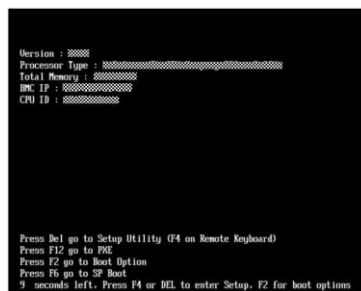
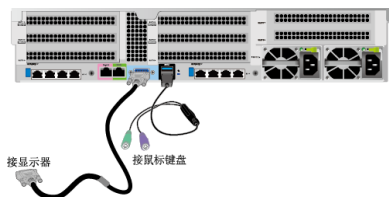
- 《Atlas硬件产品 用户清单》:

<https://support.huawei.com/enterprise/zh/doc/EDOC1100235027/ea6111a2?idPat h=23710424|251366513|22892968|252764743>。

查询管理网口IP地址（可选）

- iBMC管理网口默认IP是192.168.2.100，但是对于已经使用过的服务器，可能网址发生了更改，此时需要在BIOS中查询管理网口的IP地址。
- 首先需要通过外接设备的方式进入BIOS界面。

Step1: 连接显示器和键盘。

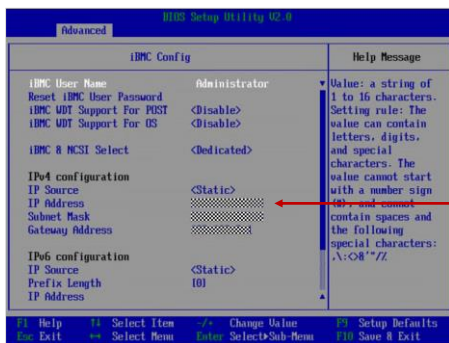


Step3: 输入密码进入BIOS界面。

- 服务器的前后面板提供了VGA接口用于连接显示器；但未提供标准的PS2键盘、鼠标接口。可以直接连接USB的键盘和鼠标，连接方法同一般的USB线缆；或通过USB转PS2线缆连接键盘和鼠标。

查询管理网口IP地址（可选）

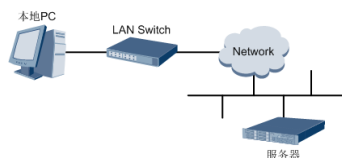
- **Step4:** 在BIOS界面查询当前管理网口IP地址。



进入BIOS界面后，选择“Advanced > IPMI iBMC Configuration> iBMC Configuration”，按“Enter”。进入“iBMC Config”界面，可以查看到IP地址。

登录iBMC Web界面

- 使用网线（交叉网线或双绞线）连接本地PC和服务器的iBMC管理网口。连接组网图如图所示。



- 打开本地PC中的浏览器，在地址栏中访问iBMC系统的地址：<https://服务器iBMC管理网口的IP地址>例如“<https://192.168.2.100>”。
- 在iBMC登录界面中，输入登录iBMC系统的用户名和密码，例如“Administrator”及“Admin@9000”。
- 在“域名”下拉列表框中，选择“这台iBMC”。单击“登录”。进入“首页”界面。
- 注意：此时可以进入“网络配置”，找到“网络协议”栏，根据实际修改iBMC管理网口地址。

欢迎到访

iBMC

用户名
请输入用户名

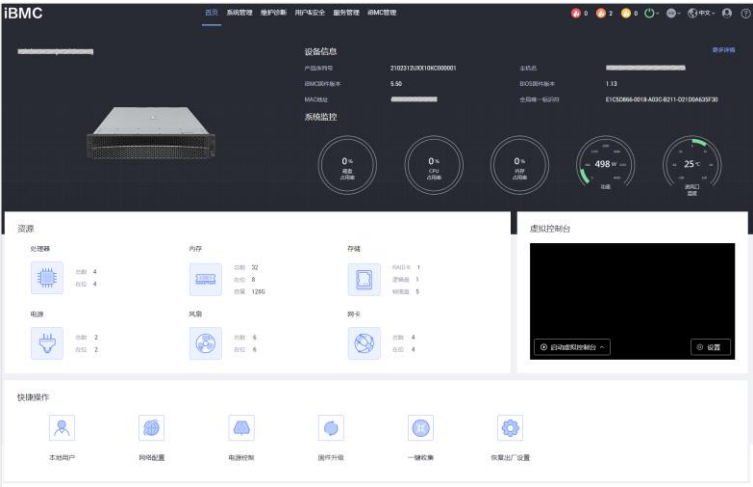
密码
请输入密码

域名
这台iBMC

登录

[单点登录](#)

iBMC首页



修改iBMC管理网口地址（可选）

- 进入iBMC首页界面，选择“iBMC管理”，进入“网络配置”。
- 找到“网络协议”栏，根据实际进行网络配置。
- 配置完成后，点击“保存”，即能配置好iBMC管理网口IP地址。

网络协议

☐ IPv4 ☐ IPv6 ☒ IPv4/IPv6

IPv4

☐ 自动获取 ☒ 手动配置

IP地址

掩码

默认网关

MAC地址

IPv6

☐ 自动获取 ☒ 手动配置

IP地址

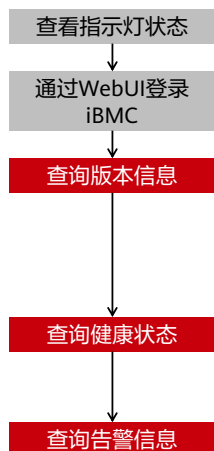
前缀长度

默认网关

链路本地地址

IP地址2

检查服务器



a.选择“iBMC管理”，在左侧导航树中选择“固件升级”查询服务器版本信息，确认服务器的版本满足局点要求。



b.在iBMC首页右上角查询服务器的健康状态。



c.若查询结果显示有告警，请参见《Atlas 800推理服务器iBMC告警处理（型号 3000）》清除产生的告警。

- 在前一步我们已经通过WebUI登录了服务器，因此这里只介绍通过WebUI检查服务器的方法。如果想要通过iBMC命令行检查服务器，可查看用户指南：
<https://support.huawei.com/enterprise/zh/doc/EDOC1100114913/3b0811f9>。
- 《Atlas 800推理服务器iBMC告警处理（型号 3000）》：
<https://support.huawei.com/enterprise/zh/ascend-computing/a800-3000-pid-250743608?category=troubleshooting&subcategory=alarm-handling>。

固件升级（可选）

- 单击固件升级区域框“...”的并选择待上传的文件。
- （可选）勾选“升级完成后立即自动重启iBMC，使升级的固件生效”前方的复选框。本选项仅适用于iBMC升级操作。
- 单击“开始升级”，并确认。
- 升级成功后，iBMC将进入自动重启并跳转至登录页面。

固件版本信息

重置iBMC

可用分区镜像切换

iBMC主用分区镜像版本	3.01.00.00
iBMC备用分区镜像版本	3.01.00.00
iBMC可用分区镜像版本	3.01.00.00
BIOS版本	1.31
CPLD版本	0.05

固件升级

SD卡控制固件升级完成之后，iBMC会自动重启以使升级的固件生效。如果在系统上电状态时升级BIOS或CPLD，则BIOS在系统下电再上电或重启后生效，CPLD在系统下电后生效。

☒ 升级完成后立即自动重启iBMC，使升级的固件生效 本选项仅适用于iBMC固件升级操作

...

开始升级

修改初始密码

- 在iBMC的Web主界面中选择“用户&安全 > 本地用户”。进入“本地用户”界面。
- 单击待修改密码的用户名后面的“编辑”。
- 按照界面信息修改用户密码。

添加				
用户ID	用户名	角色	登录接口	操作
▼ 2	Administrator	管理员	SNMP SSH IPMI Local SFTP Web Redfish	编辑 禁用 删除

配置时区&NTP

- 在导航栏中选择“iBMC管理 > 时区&NTP”界面可配置时区。当选择“DHCPv4自动获取”NTP信息时，不需要设置时区信息。

时区

地区: 其他 时区: UTC

NTP功能

NTP功能: ☒ 开启 ☐ 关闭

NTP服务器信息

☒ DHCPv4 自动获取 ☐ DHCPv6 自动获取 ☒ 手动配置

服务器一 (必选) 服务器一 (必选)

服务器二 (必选) 服务器二 (必选)

服务器三 (必选) 服务器三 (必选)

最小轮询间隔: 6 (15948s)

最大轮询间隔: 10 (172948s)

服务器身份认证: ☐ 开启 ☒ 关闭

上传NTP配置: [↓ 上传](#)

- NTP (Network Time Protocol, 网络时间协议) 就是用来使网络中的各个主机时钟同步的一种协议, 他把主机的时钟同步到协调世界时UTC, 其精度在LAN网络内可达1毫秒内, 在WAN网络上可以达到几十毫秒内。

VNC功能

- 从远端控制服务器实时桌面时需要用到的VNC（Virtual Network Console，虚拟网络控制台）服务，开启后可用本地（即用户操作所用的客户端）的鼠标、键盘对服务器进行操作，可用本地显示器查看服务器。
- 在“服务管理”的“VNC”页面，您可以查看和设置VNC功能的开启情况及相关配置项目。

VNC功能

VNC功能 ☐ 开启 ☒ 关闭

端口 [恢复默认值](#)

超时时长(分钟)

键盘布局

VNC密码

确认VNC密码

密码有效期(天) 无限制

登录规则 ☐ 规则1 允许时间： -- 至 -- 允许IP段： -- 允许MAC段： --
☐ 规则2 允许时间： -- 至 -- 允许IP段： -- 允许MAC段： --
☐ 规则3 允许时间： -- 至 -- 允许IP段： -- 允许MAC段： --
[点击跳转至“安全配置”页面修改登录规则](#)

SSL加密 ☒ 开启 ☐ 关闭

最大会话 5

活跃会话 0

SNMP功能

- SNMP（Simple Network Management Protocol，简单网络管理协议），由一组网络管理的标准组成，包含一个应用层协议、数据库模型和一组资源对象。该协议支持网络管理系统，用以监测连接到网络上的设备。
- 在“服务管理”的“SNMP”页面，您可以查看和设置SNMP功能的开启情况及相关配置项目。
- iBMC支持多个版本的SNMP：
 - SNMPv1：简单网络管理协议的第一个正式版本，在RFC1157中定义。该版本由于自身机制而存在安全隐患，请尽量避免使用。建议使用SNMPv3版本的SNMP服务。
 - SNMPv2：基于共同体的管理架构，在RFC1901中定义的一个实验性协议。该版本由于自身机制而存在安全隐患，请尽量避免使用。
 - SNMPv3：简单网络管理协议的第三个正式版本。在前面的版本基础上，SNMPv3增加了安全能力和远程配置能力。建议使用SNMPv3版本的SNMP服务。

SNMP功能

SNMP地址

☒ 开启

☐ 关闭

端口

服务器认证

联系人

位置

☐ SNMPv1

☐ SNMPv2

接口设备

☒ 开启

☐ 关闭

☐ 删除只读团体名

只读团体名

确认只读团体名

☐ 删除读写团体名

读写团体名

确认读写团体名

登录列表

☐ 编辑

名称/时间

— 是 —

名称/时

— 无2M42B —

☐ 编辑

名称/时间

— 是 —

名称/时

— 无2M42B —

☐ 编辑

名称/时间

— 是 —

名称/时

— 无2M42B —

☐ 编辑

名称/时间

— 是 —

名称/时

— 无2M42B —

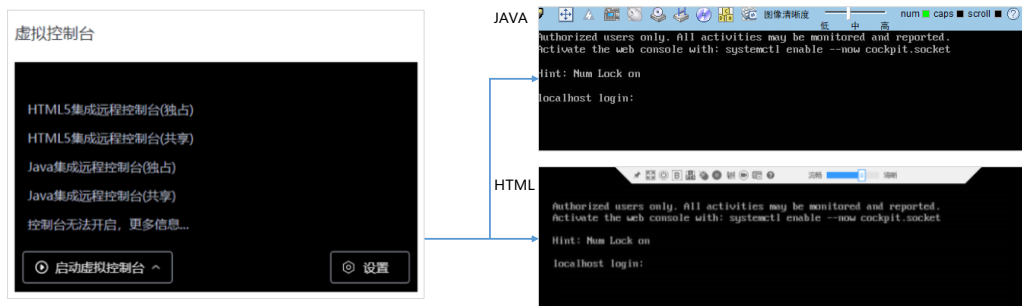
点选列表是“安全策略”页表格在登录列表

目录

1. 常见的机房结构
2. 安装部署要求
- 3. 部署关键步骤及配置项**
 - 服务器初始化配置
 - BIOS关键参数及配置
 - RAID配置
 - 操作系统安装
 - NPU驱动安装
4. 业务部署

通过iBMC WEB登录服务器远程虚拟控制台

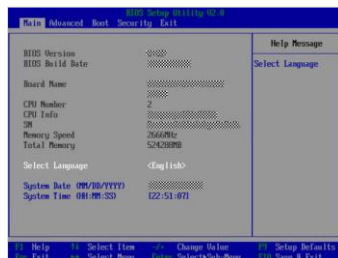
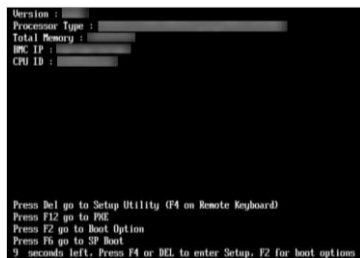
- 在“首页”右下角选择“虚拟控制台”。
- 单击“启动虚拟控制台”右侧的“^”，选择“Java集成远程虚拟控制台（独占）”、“Java集成远程虚拟控制台（共享）”、“HTML5集成远程控制台（独占）”或“HTML5集成远程控制台（共享）”，进入服务器的实时操作控制台。



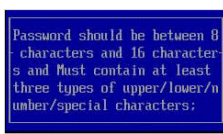
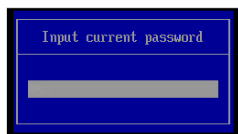
- Java集成远程虚拟控制台（独占）：只能有1个本地用户或VNC用户通过iBMC连接到服务器操作系统。
- Java集成远程虚拟控制台（共享）：可以让2个本地用户或5个VNC用户同时通过iBMC连接到服务器操作系统，并同时服务器进行操作。本用户可以看到对方用户的操作，对方用户也能看到本用户的操作。
- HTML5集成远程控制台（独占）：只能有1个本地用户或VNC用户通过iBMC连接到服务器操作系统。
- HTML5集成远程控制台（共享）：可以让2个本地用户或5个VNC用户同时通过iBMC连接到服务器操作系统，并同时服务器进行操作。本用户可以看到对方用户的操作，对方用户也能看到本用户的操作。

进入BIOS

- 在远程虚拟控制台菜单栏中，单击  或点击 ，选择“上电”或“强制重启”，将服务器上电。按“Delete”或“F4”。

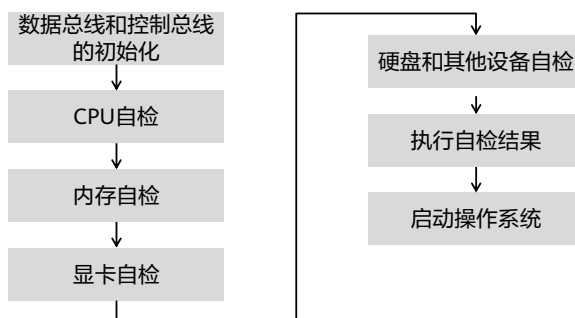


- 在弹出的“Input current password”对话框中输入当前已有的密码。若弹出设置新密码提示框时，设置并输入新密码。



上电自检

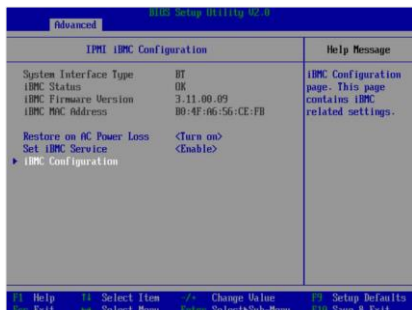
- 在进入BIOS界面的过程中，服务器会先POST（Power On Self Test，上电自检）又称为开机自检。在开机时会针对计算机硬件如CPU、主板、存储器等进行检测，结果会显示在屏幕上。
- BIOS POST阶段可以分为以下几个步骤：



- 数据总线和控制总线的初始化：BIOS首先会初始化计算机的数据总线和控制总线，以确保数据可以在各个硬件设备之间正常传输。
- CPU自检：BIOS会对处理器进行自检，检查它的型号、主频和功能是否正常。同时，BIOS还会检查处理器的缓存是否可用，并设置相应的工作模式。
- 内存自检：BIOS会对系统中的内存进行自检，以确保内存模块安装正确且可用。在这个过程中，BIOS会检查内存的容量和速度，并设置内存的工作模式和时钟频率。
- 显卡自检：BIOS会检测并初始化显卡，以确保显示设备可以正常工作。这个过程中，BIOS会检查显卡的型号和显存容量，并设置相应的显示模式。
- 硬盘和其他设备的自检：BIOS会逐个检测硬盘、光驱、USB设备等外部设备，以确保它们可以被正常识别和使用。
- 执行自检结果：BIOS会将自检的结果显示在屏幕上，通常以一系列的短信号（beep）或者显示信息的方式呈现。这些信息可以帮助用户诊断硬件问题。
- 启动操作系统：最后，BIOS会尝试从硬盘或其他可启动设备上加载操作系统，并将控制权交给操作系统。

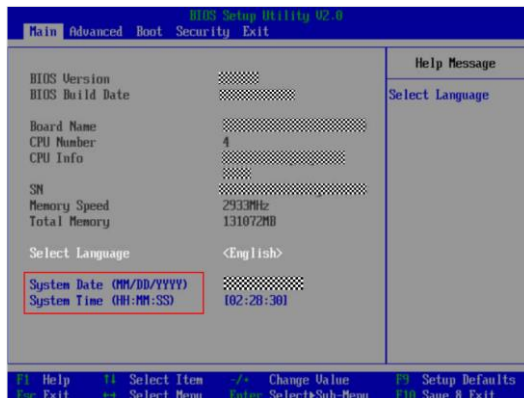
iBMC配置

- 进入“Advanced”页签，按向下键将光标移动到“IPMI iBMC Configuration”项，然后按回车，进入iBMC配置页。
- 按向下键，将光标移动到“iBMC Configuration”项，按回车键进入配置。
- 在此页中可以对iBMC的登录信息进行配置，包括登录的用户名、密码、IP地址等信息。



设置系统日期和时间

- 按“←”、“→”方向键切换至“Main”界面，选择“System Date”设置系统日期，选择“System Time”设置系统时间。



设置服务器启动方式顺序

- 按“←”、“→”方向键切换至“Boot”界面。选择“Boot Type Order”。
- 选择要设置的启动项，按“+”、“-”向上或向下移动启动项，调整启动项的排列顺序。设置完成后，按“F10”。弹出“Save configuration changes and exit?”对话框。选择“Yes”并按“Enter”保存设置。服务器将自动重启使设置生效。



- PXE（预启动执行环境）是由Inter公司开发的网络引导技术，工作在Client/Server模式，允许客户机通过网络从远程服务器下载引导镜像，并加载安装文件或整个操作系统。优势在于不需要光盘、U盘等安装介质，且可以远程同时装配多台服务器。

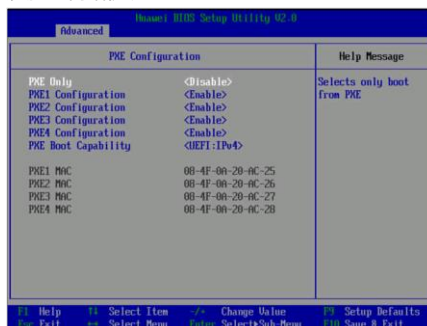
设置网卡的PXE（可选）

- 如果需要通过UEFI方式启动，需要设置网卡的PXE，以设置板载网卡的PXE为例。
- 切换至“Advanced”页签，选择“LOM Configuration>PXE Configuration”，进入PXE Configuration界面。
- 选择要配置的网口如“PXE1 Configuration”，在弹出的菜单选项对话框中选择“Enable”开启对应网口的PXE功能。
- 选择“PXE Boot Capability”，在弹出的菜单选项对话框中选择需要支持的网络协议。

- UEFI: IPv4
- UEFI: IPv6
- UEFI: IPv4/IPv6

- 设置完成后，按“F10”，选择“Yes”并按“Enter”保存设置。

服务器将自动重启使设置生效。



- 设置PCIe网卡的PXE见用户指南：

<https://support.huawei.com/enterprise/zh/doc/EDOC1100114913/65834e2e>。

iBMC WEB配置BIOS关键参数

- 通过“BIOS配置”功能，设置操作系统第一启动项。
- 在导航栏中选择“系统管理 > BIOS配置”，选择优先引导介质，支持仅在下次重启时生效或永久生效。

The screenshot displays the BIOS configuration interface for iBMC. It features a 'Boot Order' (启动顺序) section with a dropdown menu for 'Priority Boot Device' (优先引导介质) set to 'Hard Disk' (硬盘). Below this, a list of boot devices is shown: 'Hard Disk Device' (硬盘设备), 'Optical Drive' (光盘装置), 'PXE', and 'Other' (其他). The 'Hard Disk Device' is currently selected. To the right of the list, there are two radio buttons: 'Single Effective' (单次有效) and 'Permanent Effective' (永久有效), with the latter being selected. A 'Save' (保存) button is located at the bottom left of the configuration area.

目录

1. 常见的机房结构
2. 安装部署要求
- 3. 部署关键步骤及配置项**
 - 服务器初始化配置
 - BIOS关键参数及配置
 - **RAID配置**
 - 操作系统安装
 - NPU驱动安装
4. 业务部署

查看服务器RAID信息

- 登录iBMC WebUI，在上方导航栏中选择“系统管理 > 系统信息”，打开“系统信息”界面。
- 在“其他”页签中查看RAID控制卡的型号信息。

The screenshot shows the iBMC WebUI interface. At the top, there are tabs for '产品信息' (Product Information), '处理器' (Processor), '内存' (Memory), '网卡' (Network Card), '传感器' (Sensor), and '其他' (Other). The '其他' tab is selected. On the left, there is a sidebar with '硬盘阵列 (1/1)' (Disk Array (1/1)) expanded, showing 'RAID卡 (1/1)' (RAID Card (1/1)) selected. The main area displays 'RAID卡 (1/1)' (RAID Card (1/1)) with a table of RAID card information. The table has columns: 名称 (Name), 位置 (Location), 厂商 (Manufacturer), 编号 (ID), 型号 (Model), PCB版本 (PCB Version), CPLD版本 (CPLD Version), 部件编码 (Part Number), 序列号 (Serial Number), 单板ID (Board ID), and 资源归属 (Resource Ownership). The first row shows 'SR450C-M 20' (Name), 'mainboard' (Location), 'Huawei' (Manufacturer), '1' (ID), 'LSI SAS3508' (Model, highlighted with a red box), 'B' (PCB Version), '0.02' (CPLD Version), '03024JMY' (Part Number), '033EFTT10KA001749' (Serial Number), '0x002a' (Board ID), and 'CPU1' (Resource Ownership).

名称	位置	厂商	编号	型号	PCB版本	CPLD版本	部件编码	序列号	单板ID	资源归属
SR450C-M 20	mainboard	Huawei	1	LSI SAS3508	B	0.02	03024JMY	033EFTT10KA001749	0x002a	CPU1

配置RAID (9460-8i)

- 以Avago MegaRAID SAS 9460-8i为例，以下为该产品支持的RAID类型及对应的硬盘个数：

RAID级别	支持硬盘数	支持子组数	子组支持硬盘数	允许坏盘数
RAID 0	1 ~ 32	不涉及	不涉及	0
RAID 1	2~32 (偶数)			硬盘数 ÷ 2
RAID 5	3 ~ 32			1
RAID 6	3 ~ 32			2
RAID 10	4~240 (偶数)	2 ~ 8个RAID 1	2~32 (偶数)	子组数
RAID 50	6 ~ 240	2 ~ 8个RAID 5	3 ~ 32	子组数
RAID 60	6 ~ 240	2 ~ 8个RAID 6	3 ~ 32	子组数x2

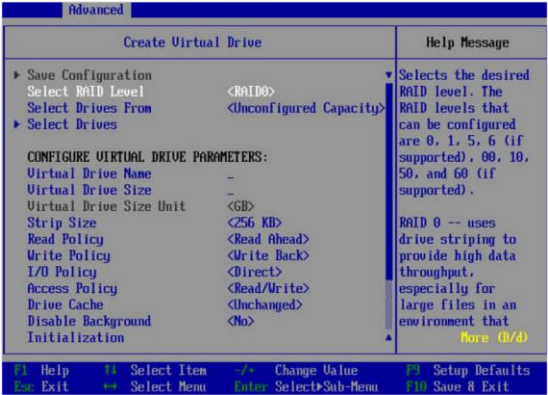
- 华为系列RAID卡配置文档：
<https://support.huawei.com/enterprise/zh/doc/EDOC1100048779?idPath=23710424|251366513|22892968|252309139|250743608>。
- 不支持RAID 00级别。
- 子组数：子RAID的个数，例如RAID 50由两个RAID 5组成，则子组数为2。
- RAID 10、RAID 50每个子组中最多允许1个坏盘。
- RAID 60每个子组中最多允许2个坏盘。
- RAID10、50、60包含的总硬盘数由子组数和子组支持的硬盘数共同决定。

JBOD硬盘直通

- 硬盘直通，即“JBOD”功能，又称指令透传，是不经过传输设备处理，仅保证传输质量的一种数据传输方式。
- 打开硬盘直通功能后，RAID控制卡可对所连接的硬盘进行指令透传，在不配置RAID组的情况下，用户指令可以直接透传到硬盘，方便上层业务软件或管理软件访问控制硬盘。
- 例如，服务器操作系统安装过程中，可以直接找到挂载在RAID控制卡下的硬盘作为安装盘；而不支持硬盘直通的RAID控制卡，在操作系统安装过程中，只能找到该RAID控制卡下已经配置好的虚拟磁盘作为安装盘。

VNC配置RAID (1)

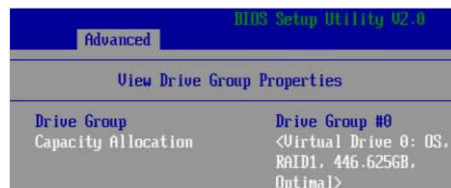
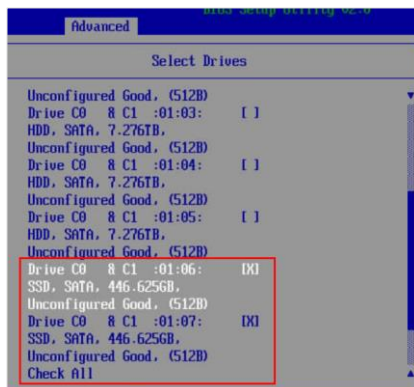
- 在BIOS配置的“Advanced”页签中找到AVAGO RAID卡配置项。依次进入“Main Menu”、“Configuration Management”、“Create Virtual Drive”，进入RAID创建页面。



参数	参数说明
Select RAID Level	选择RAID级别。
Select Drives From	虚拟磁盘成员盘来源。 成员盘来源分为： •Unconfigured Capacity: 未加入过虚拟磁盘的空闲硬盘。 •Free Capacity: 硬盘组中未作为虚拟磁盘使用的空间。
Select Drives	选择成员盘。
Access Policy	数据读取的策略，分以下三种： •Read/Write: 可读可写。 •Read Only: 只读访问。 •Blocked: 禁止访问。

VNC配置RAID (2)

- 在“Select Drives”中选择两块SSD磁盘，按空格选中硬盘，然后找到最下方的“Apply Changes”确认。
- 在“View Drive Group Properties”中，可以看到有一个已经配置好的RAID1。



iBMC WEB配置RAID (1)

- 在iBMC的web界面，点击“系统管理”-->“存储管理”，进入iBMC的RAID卡配置界面。依次展开“PCIe Card”、“Logical Drive”，可以看到我们在BIOS中创建的RAID1以及其用到的两块磁盘。

The screenshot shows the iBMC web interface. The top navigation bar includes 'System Management' (系统管理), 'Maintenance' (维护诊断), 'User & Security' (用户&安全), 'Service Management' (服务管理), and 'iBMC Management' (iBMC管理). The left sidebar shows 'System Management' (系统管理) with sub-items like 'System Information' (系统信息), 'Performance Monitoring' (性能监控), 'Storage Management' (存储管理), 'Power & Power' (电源&功率), 'Fan & Cooling' (风扇&散热), and 'BIOS Configuration' (BIOS配置). The 'Storage Management' (存储管理) section is active, showing a tree view of storage components. Under 'PCIe Card 3 (9450-8)', 'Logical Drive-0' is selected. The main area displays RAID configuration details for RAID1, including disk selection, stripe size, and write policy.

名称	OS	RAID级别	RAID1
条带大小	256 KB	状态	Optimal
容量	446.625 GB	启动设备	是
当前写策略	Write Back With BBU	默认写策略	Write Back With BBU
用于二倍缓存	否	初始化类型	No Init
一致性校验	已禁用	默认读策略	Read Ahead
默认IO策略	Direct IO	后台初始化	已启用
当前读策略	Read Ahead	当前IO策略	Direct IO
读写策略	Read Write	物理盘缓存状态	Disk's Default

iBMC WEB配置RAID（2）

- RAID配置也可以通过iBMC的web界面进行配置。点击“PCIe Card”后面的“添加”。在弹出的“创建逻辑盘”窗口中配置新RAID的信息。

The image shows two screenshots from the iBMC Web interface. The left screenshot shows the 'Storage Management' (存储管理) page with a tree view containing 'PCIe Card 3 (9460-S)' and 'Logical Drive 0'. A red box highlights the 'Add' (添加) button next to the PCIe Card. The right screenshot shows the 'Create Logical Disk' (创建逻辑盘) dialog box. In this dialog, several fields are highlighted with red boxes: the name 'Data', the initialization type 'Quick Init', the RAID level '5', and the physical disks 'Disk0', 'Disk1', 'Disk2', 'Disk3', and 'Disk4' which are all selected with checkboxes.

- 名称设置为“Data”；
- 初始化类型设置为“Quick Init”；
- RAID级别设置为5；
- 服务器中安装4块8 T容量的SATA HDD磁盘（磁盘编号可以从存储管理处依次点击每个磁盘查看磁盘信息，8 T容量的实际可用容量大约为7.2 T）。

目录

1. 常见的机房结构
2. 安装部署要求
- 3. 部署关键步骤及配置项**
 - 服务器初始化配置
 - BIOS关键参数及配置
 - RAID配置
 - 操作系统安装
 - NPU驱动安装
4. 业务部署

服务器兼容性的考量

- 在前一章我们知道操作系统是一种控制计算机的特殊计算机程序。操作系统可以帮助用户建立与服务器的连接，且能在多用户使用的情况下协调分配有限资源。比如CPU、磁盘、内存和打印机等，用户运行程序时需要访问这些资源。
- 由于不同的操作系统有不同版本和不同的配置，在计算设备设计之初，只对某些特定的操作系统进行兼容性测试，确保操作系统可以正常运行。
- 因此在安装服务器操作系统前，需要先查明服务器的操作系统兼容性和部件兼容性信息。

Atlas 800 - 3000兼容性

- 以Atlas 800 推理服务器（型号 3000）为例，该服务器可以支持CentOS、Ubuntu和openEuler等部分版本的操作系统。
- 如果需要查询服务器最新的操作系统兼容性和部件兼容性信息，可以通过访问昇腾计算兼容性助手查询。

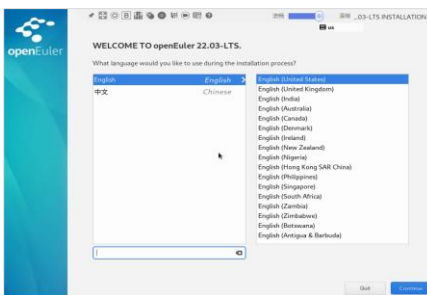
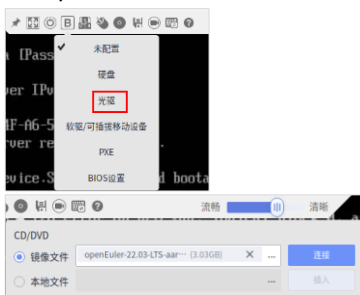
操作名称	特性描述
CentOS 7.6 for ARM	CentOS Linux 7 Update 6 Server for ARM
CentOS 8.2 for ARM	CentOS Linux 8 Update 2 Server for ARM
Kylin Linux Advanced Server V10 SP1 for ARM	Kylin Linux Advanced Serve V10 SP1 for ARM
NeoKylin Server V7.0 U6 for ARM	Neokylin Server V7.0 U6 for ARM
SLES 12 SP5 for ARM	SUSE Linux Enterprise Server 12 Service Pack 5 for ARM
UOS 20 Server for ARM	UOS 20 Server for ARM
UOS20-1020e Server for ARM	UOS20-1020e Server for ARM(powered by openEuler)
Ubuntu 20.04 LTS for ARM	Ubuntu 20.04 LTS Server Edition for ARM
openEuler 20.03 LTS for ARM	openEuler 20.03 LTS for ARM (powered by openEuler)
openEuler 22.03 LTS for ARM	openEuler 22.03 LTS for ARM (powered by openEuler)



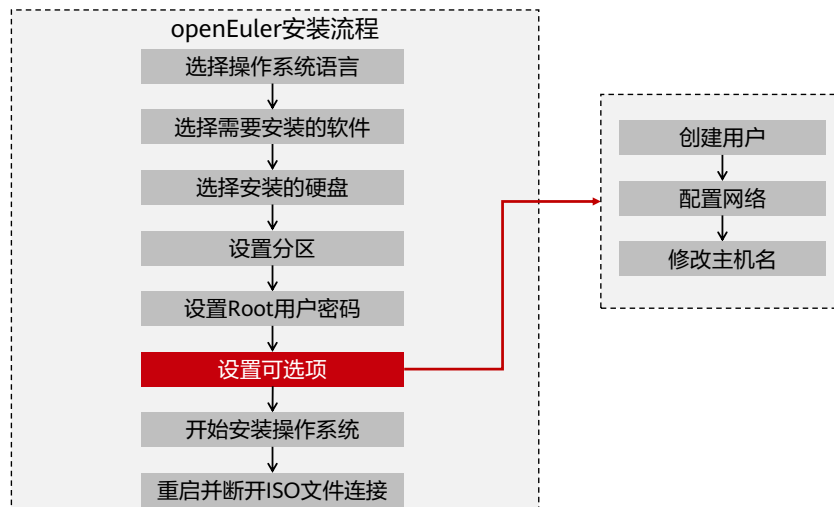
- 昇腾计算兼容性助手：
<https://info.support.huawei.com/computing/ftca/zh/product/atlas>。

挂载ISO

- 通过iBMC WEB登录服务器远程虚拟控制台。
- 在控制台将Boot启动项修改为光驱，并选择光驱，如“openEuler-22.03-LTS-aarch64-dvd.iso”。
- 重启服务器进入系统安装界面，选择“Install openEuler 22.03-LTS”选项。
- 进入openEuler安装界面，选择中文/英文语言进行安装。

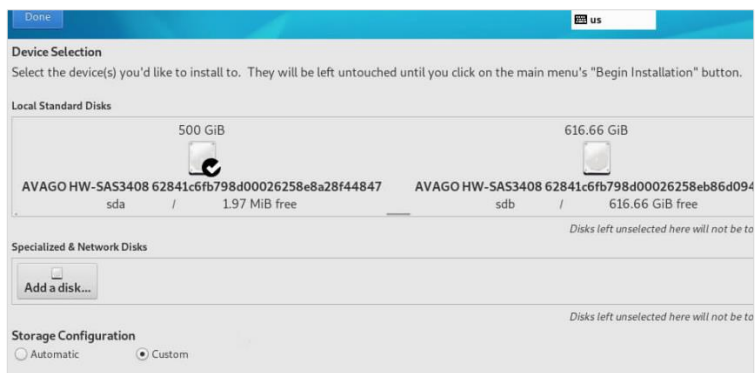


操作系统安装步骤概览



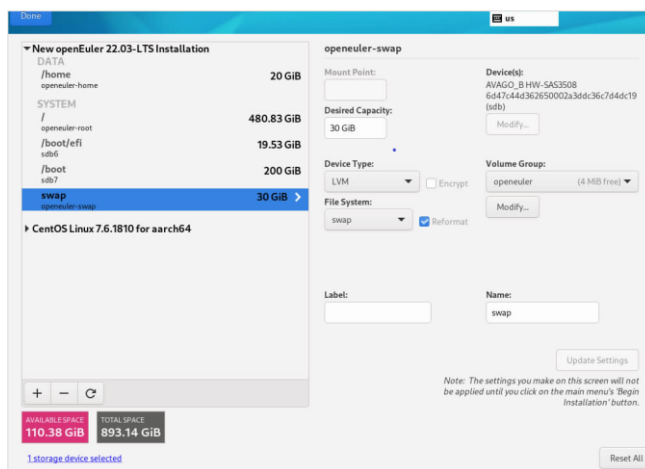
设置分区

- 在安装操作系统时，默认为自动分区。在“Storage Configuration”中选择“Custom”可以进入硬盘分区界面。



根据实际情况创建硬盘分区

- openEuler是Linux操作系统，Linux系统是以文件的形式对计算机中的数据和硬件资源进行管理的。
- 在安装过程中，用户可以为每个文件分配合适的存储空间。



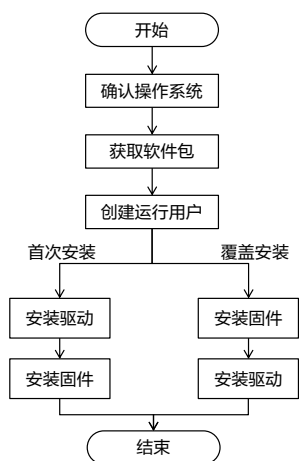
目录

1. 常见的机房结构
2. 安装部署要求
- 3. 部署关键步骤及配置项**
 - 服务器初始化配置
 - BIOS关键参数及配置
 - RAID配置
 - 操作系统安装
 - NPU驱动安装
4. 业务部署

驱动和固件的区别

- 驱动和固件都是代码，前者为软件服务，后者为硬件服务。
- 驱动程序一般指的是设备驱动程序（Device Driver），是一种可以使计算机和设备进行相互通信的特殊程序。相当于硬件的接口，操作系统只有通过这个接口，才能控制硬件设备的工作，假如某设备的驱动程序未能正确安装，便不能正常工作。因此，驱动程序被比作“硬件和系统之间的桥梁”。
- 像是打印机、显卡、网卡、声卡等都需要驱动程序帮助操作系统直接访问硬件资源。驱动软件需要提前下载在服务器上，并确保和操作系统兼容。
- 固件（Firmware）是预安装到硬件产品内部只读记忆体里面，与硬件产品捆绑匹配的程序，在硬件设备内部。此类软件具有不被修改、与硬件捆绑的特性，脱离于操作系统，不需要考虑和操作系统兼容的问题。
- 由于不同的操作系统对于操作硬件的方式完全不同，只使用固件需要有太多的版本。所以我们需要与操作系统对应的驱动帮助我们适配不同的操作系统。

NPU驱动和固件安装流程



- 驱动和固件的安装用户必须为root用户，驱动和固件的运行用户可以为root用户或者非root用户，如果后续用户对运行业务有用户权限控制或者多用户隔离的需求，可创建非root用户作为后续运行用户。
- 硬件设备前期安装过驱动固件且未卸载，属于覆盖安装场景，需要再次安装驱动固件。
- 驱动的安装有三种方式：二进制文件直接安装、源码编译安装和重构驱动包安装。
 - 如果需要自定义驱动：选择重构驱动包安装。
 - 不需要自定义驱动，在对比运行环境OS内核版本和《版本配套表》中OS版本对照后：如果不一致则需要选择源码编译安装，如果一致则源码编译和二进制安装都可以。

目录

1. 常见的机房结构
2. 安装部署要求
3. 部署关键步骤及配置项
- 4. 业务部署**
 - openEuler操作
 - openGauss操作

FTP服务器

- FTP（File Transfer Protocol）即文件传输协议，是互联网最早的传输协议之一，其最主要的功能是服务器和客户端之间的文件传输。FTP使用户可以通过一套标准的命令访问远程系统上的文件，而不需要直接登录远程系统。另外，FTP服务器还提供了如下主要功能：
 - 用户分类：默认情况下，FTP服务器依据登录情况，将用户分为实体用户（real user）、访客（guest）、匿名用户（anonymous）三类。三类用户对系统的访问权限差异较大，实体用户具有较完整的访问权限，匿名用户仅有下载资源的权限。
 - 命令记录和日志文件记录：FTP可以利用系统的syslogd记录数据，这些数据包括用户历史使用命令与用户传输数据（传输时间、文件大小等），用户可以在/var/log/中获得各项日志信息。
 - 限制用户的访问范围：FTP可以将用户的工作范围限定在用户主目录。用户通过FTP登录后系统显示的根目录就是用户主目录，这种环境被称为change root，简称chroot。这种方式可以限制用户只能访问主目录，而不允许访问/etc、/home、/usr/local等系统的重要目录，从而保护系统，使系统更安全。

守护进程vsftpd

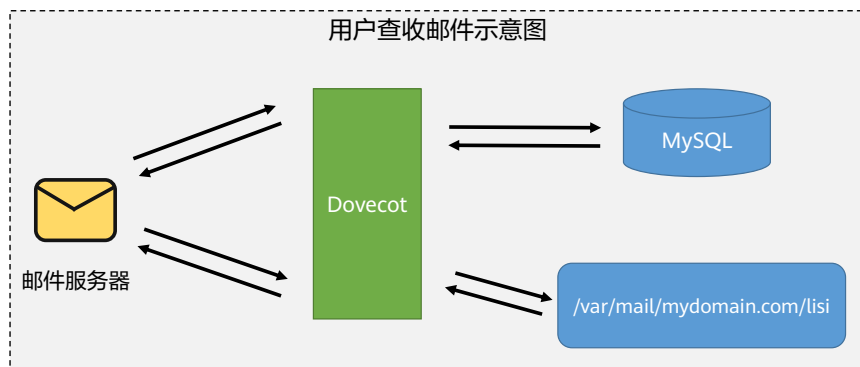
- 由于FTP历史悠久，它采用未加密的传输方式，所以被认为是一种不安全的协议。为了更安全地使用FTP，这里介绍FTP较为安全的守护进程vsftpd（Very Secure FTP Daemon）。
- 之所以说vsftpd安全，是因为它最初的发展理念就是构建一个以安全为中心的FTP服务器。它具有如下特点：
 - vsftpd服务的启动身份为一般用户，具有较低的系统权限。此外，vsftpd使用chroot改变根目录，不会误用系统工具。
 - 任何需要较高执行权限的vsftpd命令均由一个特殊的上层程序控制，该上层程序的权限较低，以不影响系统本身为准。
 - vsftpd整合了大部分FTP会使用到的额外命令（例如dir、ls、cd等），一般不需要系统提供额外命令，对系统来说比较安全。

基于FTP的私有yum源部署流程

- Step 1: 在本地新增yum源。
- Step 2: 安装FTP服务，并设置FTP开机自启动。
- Step 3: 修改vsftp配置文件，允许匿名用户访问，并关闭写权限。
- Step 4: 关闭防火墙和selinux。
- Step 5: 设置自动挂载。
- Step 6: 验证FTP服务。

在Linux系统上搭建邮箱服务器

- 本方案以主流的postfix+dovecot为基础，其中postfix用作邮件传输协议smtp，dovecot用作pop3（或imap）服务器。



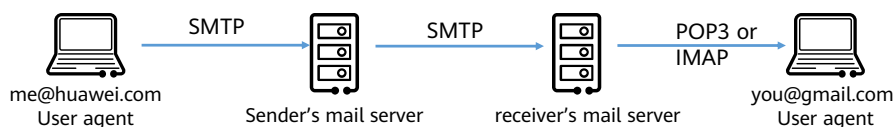
- Postfix：一个开源的电子邮件服务器。
- SMTP简单邮件传输协议定义了一个在电子邮件客户端和邮件服务器之间交换数据的过程。
- Dovecot：CentOS系统中著名的POP3/IMAP服务器实现。POP3/IMAP是从邮件服务器中读取邮件时使用的协议，POP3将邮件全部从服务器中拖取下来；IMAP则每次从服务器获取邮件名等关键信息，要读某封邮件时，才从服务器下载。

邮箱服务器的工作模式

- 用postfix构建的邮件系统至少有两种工作模式：
 - 第一种：利用本地Linux账号进行邮件收发，比如本地系统有用户root或someone，那么就有root@example.com和someone@example.com两个email地址。
 - 第二种：为了管理的方便和系统安全，postfix的用户管理采用虚拟用户方式，即postfix单独设立了许多用户，他们各自在系统中映射有独立的硬盘空间。但同时这些用户又跟本地Linux系统内固有的真实账号没有关联。这种方式相对第一种要更复杂，适合于多用户的场景。
- 在服务器中搭建一个邮件系统，如果直接在linux中创建该系统，邮箱服务器用户数会收到uid数量的限制。所以需要用第二种虚拟用户的方法来解决这个问题。

邮箱服务器运行原理

- 发件人me@huawei.com向收件人you@gmail.com的电子邮件，在smtp协议下的投递流程：
 - me@huawei.com用邮件客户端（比如outlook）写了一封邮件给你@ gmail.com，点下发送按钮后，邮件首先会发送到smtp.huawei.com。
 - smtp.huawei.com检索到这封邮件的收件人域名是gmail.com，于是通过互联网将邮件发送到smtp.gmail.com。
 - smtp.gmail.com确认收下邮件后，将它转存到邮件服务器的硬盘中待收。
 - 此过程中smtp服务器起到了“收、发”的作用。对于smtp.huawei.com而言是发送，对于smtp.gmail.com则是接收。
- Pop3（或imap）收发服务器在方案中起到中转作用，将存储在邮件服务器硬盘中的邮件转移回邮件客户端（user agent），只负责从邮件服务器到邮件客户端这段路径，而邮件收发则由smtp完成。



- pop3与imap两种方式的区别是，pop3将邮件拉回本地后，即脱离服务器。imap则更先进些，它能实时将你在邮件客户端的操作反馈回邮件服务器，比如：删除邮件，标记已读等，服务器上的邮件也会做相应的动作。所以无论从浏览器登录邮箱或者客户端软件登录邮箱，看到的邮件以及状态都是一致的。

目录

1. 常见的机房结构
2. 安装部署要求
3. 部署关键步骤及配置项
- 4. 业务部署**
 - openEuler操作
 - openGauss操作

openGauss部署方式

- openGauss支持单机部署和单机HA部署两种部署方式。单机部署时，可在一个主机部署多个数据库实例，但为了数据安全，不建议用户这样部署。单机HA部署支持一台主机和最少一台备机，备机一共最多8台的配置方式。
- 通过openGauss提供的脚本安装时，只允许在单台物理机部署一个数据库系统。如果您需要在单台物理机部署多个数据库系统，建议您通过命令行安装，不需要通过openGauss提供的安装脚本执行安装。

openGauss安装部署

- 安装前准备：openGauss安装前需要准备相应的软硬件环境以及相关配置。
- 获取并校验安装包：安装包需要在官方网站上下载并且对安装包进行校验。
- 配置XML文件：安装openGauss前需要创建XML文件。XML文件包含部署openGauss的服务器信息、安装路径、IP地址以及端口号等。
- 初始化安装环境：安装环境的初始化包含上传安装包和XML文件、解压安装包、使用gs_preinstall准备好安装环境。
- 执行安装：使用gs_install安装openGauss。
- 设置备机可读：可选操作。在开启备机可读之后，备机将支持读操作，并满足数据一致性要求。



思考题

1. （单选题）如果想要以“openEuler-22.03-LTS-aarch64-dvd.iso”文件来安装操作系统，需要将Boot启动项修改为以下哪一项？
- A. 硬盘
 - B. 光驱
 - C. 软驱
 - D. PXE

- 参考答案：B

本章总结

- 本章我们一起学习了常见的机房结构
- 熟悉了服务器安装部署的要求，部署过程中的关键步骤及配置方式
- 了解了如何在服务器上部署业务

缩略语

- NPU: Neural Processing Unit, 神经处理单元, 是一类专用于人工智能硬件加速的微处理器或计算系统。
- BIOS: Basic Input/Output System, 基本输入输出系统, 是固化到计算机中的一组程序, 为计算机系统提供最底层的、最直接的硬件控制服务, 解决硬件的实时需求。
- RAID: Redundant Array of Independent Disks, 独立硬盘冗余阵列。
- DHCP: Dynamic Host Configuration Protocol, 动态主机设定协议, 是一个用于IP网络的网络协议。
- POST: Power On Self Test, 上电自检。指计算机系统, 接通电源, (BIOS程序) 的行为, 包括对CPU、系统主板、基本内存、扩展内存、系统ROM BIOS等器件的测试。
- PXE: Preboot eXecution Environment, 预启动执行环境, 提供了一种使用网络接口 (Network Interface) 启动计算机的机制。
- UEFI: Unified Extensible Firmware Interface, 统一可扩展固件接口, 用来定义操作系统与系统固件之间的软件界面, 作为BIOS的替代方案。

缩略语

- JBOD: Just a Bunch of Disks, 磁盘簇, 指还没有根据RAID系统配置以增加容错率和改进数据访问性能的电脑硬盘。
- SMTP: Simple Mail Transfer Protocol, 简单邮件传输协议, 定义了电子邮件客户端和邮件服务器之间交换数据的过程。

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。
Bring digital to every person, home, and
organization for a fully connected,
intelligent world.

Copyright©2024 Huawei Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.



计算产品基础运维



前言

- 随着计算产业的发展，工程师每年负责交付和维护的设备数量成倍的增长，从最初的每年几万台到几十万台，再到现在的百万级。纯靠人工的交付和维护已经无法满足每年百万级设备量的交付和维护。面向开发者的单机管理接口、面向服务工程师轻量便捷的批量工具、面向客户的自动化集中网管、以及跨栈跨域的智能化管理软件，都是必不可少的。

目标

- 学完本课程后，您将能够：
 - 了解计算产品运维软件体系
 - 熟悉单机带内管理iBMA的功能
 - 熟悉单机带外管理iBMC、Smart Provisioning的功能
 - 掌握集中管理FusionDirector的功能
 - 了解集中管理eSight的功能
 - 掌握近端运维SmartKit的功能
 - 了解远端运维eService的功能

目录

1. 计算运维软件体系

2. 单机管理

3. 集中管理

4. 近端运维

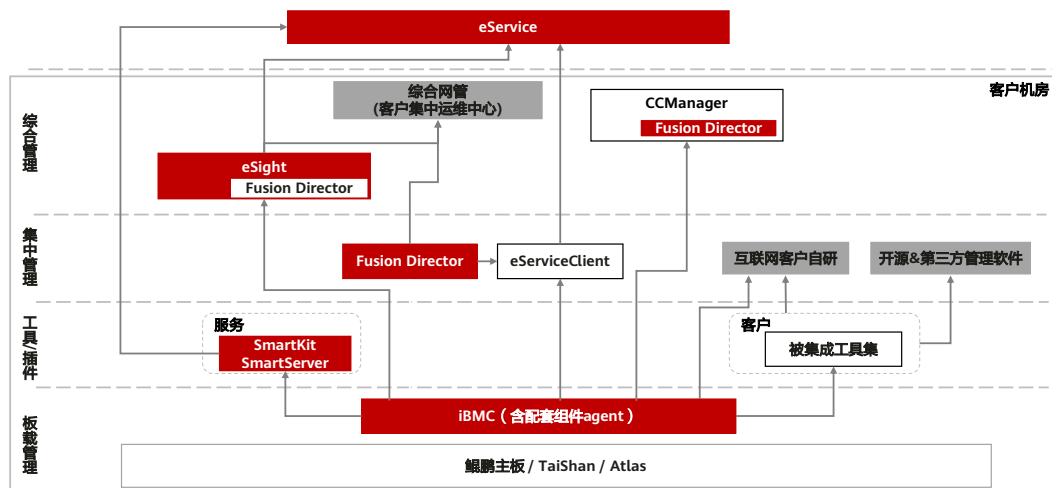
5. 远端运维

计算产品运维体系演变



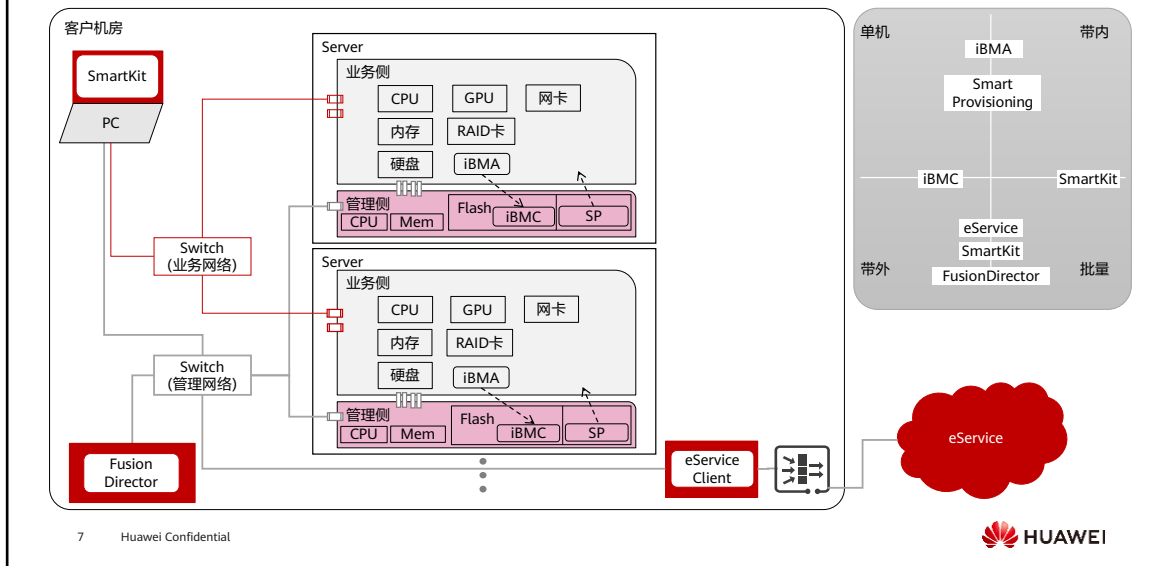
- 随着计算产业的发展，一线工程师每年负责交付和维护的设备数量成倍的增长，从最初的每年几万台到几十万台，再到现在的百万级。纯靠人工的交付和维护已经无法满足每年百万级设备量的交付和维护。工具化、自动化、智能化是运维管理软件发展的方向。是不是说我们最终的目标是智能化，直接跳过工具化、自动化就可以呢？答案显然不是，面向开发者的单机管理接口、面向服务工程师轻量便捷的批量工具、面向客户的自动化集中网管、以及跨栈跨域的智能化管理软件，都是必不可少的。

华为计算运维软件总体架构



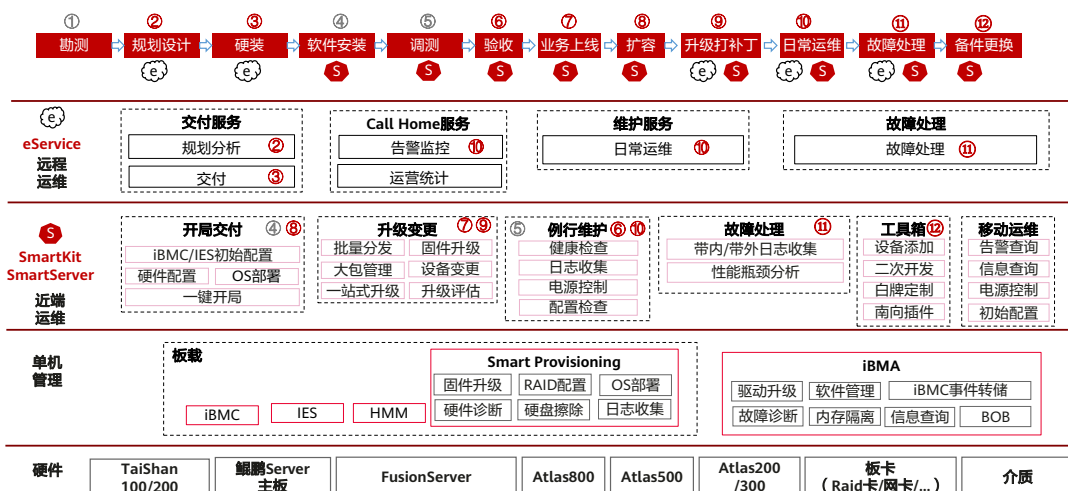
- 华为计算的产品包括FusionServer、鲲鹏、TaiShan、Atlas和FusionPod等。
- 随每台设备发货的iBMC是计算产品管理软件的基础，提供单机管理能力和北向接口。
- 对于服务工程师，我们提供SmartKit，提升近端交付和维护的效率。
- 对于互联网类开发能力较强的客户，我们提供SDK，供客户的网管集成使用。
- 对于企业类客户，我们提供集中网管FusionDirector，为客户提供设备的日常监控和管理能力。
- 对于大型客户，如运营商，我们提供eSight，提供跨域的管理能力，面向数据中心，我们提供CCManager，提供跨栈和跨域的管理能力。
- 对于长尾客户，如设备量不大，不愿意投入精力做设备运维的企业，我们提供eService，提供远程代维能力。

华为计算运维软件应用场景



- 一台服务器就如同一台高配的PC，由内存、CPU、硬盘等硬件组成。服务器为了更好的管理硬件，增加了管理面，管理面通过硬件信号监控和管理服务器上的硬件，并提供交互界面和各种管理接口IPMI/redfish/...，这就是iBMC。iBMC有自己的CPU、内存和存储资源，iBMC的运行不会影响到业务侧。iBMC通过硬件信号的管理无法满足用户的需求，就有了iBMA，iBMA可以支持服务器CPU、内存、硬盘、网卡等各个组件的状态监控、性能监控、事件监控，并把这些信息传给iBMC。服务器要跑业务（运行数据库、作为网站服务器等），就得先安装系统和驱动，服务器从iBMC的Flash上划分出了一部分空间存放无盘系统，该无盘系统用来引导客户安装操作系统和驱动，即Smart Provisioning，简称SP。
- 客户机房一般都是多台服务器，服务器之间通过交换机连接起来。SmartKit通过管理网络调各设备的iBMC接口，就可以实现对设备的批量管理，如批量配置、部署、升级、诊断。通过业务网络登录各服务器的业务侧，就可以实现对业务系统的批量管理，如批量巡检和日志收集。SmartKit定位是服务和运维工具，一般一次进站，操作完成后即把数据删除。当用户需要实时监控，对设备进行7*24不间断管理时，就需要在机房部署一个集中管理软件，这就是FusionDirector。FusionDirector提供了针对服务器的设备发现、监控、配置、部署、升级等服务器管理功能。当客户不想自己管理服务器，而希望华为能提供管理服务，就有了eService。
- 从技术的角度，网络管理可分为带外管理（out-of-band）和带内管理（in-band）两种管理模式。所谓带内管理，是指网络的管理控制信息与用户网络的承载业务信息通过同一个逻辑信道传送；所谓带外管理，是指网络的管理控制信息与用户网络的承载业务信息在不同的逻辑信道传送。

华为计算运维软件全景



- 这页图包含了服务工程师可用的工具全景图，包括SP（Smart Provisioning）、SmartKit和eService。列出了各个工具所支撑的运维场景及各个场景下提供的功能，功能很多，大家安装或使用过这些工具，就可以看到，只是各个功能使用的频率上会有差别。FD（FusionDirector）没有列在这个图中，后面有FD的小节专门为大家介绍。图中e代表eService，S代表SmartKit。
- IES：Atlas Intelligent Edge System，Atlas智能边缘管理系统。
- HMM：Hyper Management Module，超级管理模块，X6800的HMM是实现服务器管理的控制单元，兼容服务器业界管理标准IPMI2.0规范。HMM负责电源和风扇的集中管理，是服务器节点iBMC的管理延伸，主要负责电源和风扇健康状态监控、资产信息和实时数据获取、散热调速、网络汇聚。

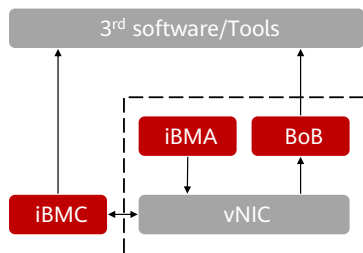
目录

1. 计算运维软件体系
2. 单机管理
 - 带内管理
 - 带外管理
3. 集中管理
4. 近端运维
5. 远端运维

- 从技术的角度，网络管理可分为带外管理（out-of-band）和带内管理（in-band）两种管理模式。所谓带内管理，是指网络的管理控制信息与用户网络的承载业务信息通过同一个逻辑信道传送；所谓带外管理，是指网络的管理控制信息与用户网络的承载业务信息在不同的逻辑信道传送。

iBMA概述

- iBMA作为带内管理代理软件，对上层管理软件iBMC开放了标准的Redfish管理接口，且可以通过BOB映射支持与带外系统的融合，以提供更完整的服务器设备管理能力。
 - 运行在带内操作系统侧，聚焦单节点的硬件设备管理，与iBMC形成互补的关系，管理iBMC无法直接管理的硬件设备。
 - iBMA为硬件和操作系统的界面，通过操作系统侧的信息获取和分析，支撑操作系统和硬件设备的定界。
 - iBMA带内获取硬件侧的配置和性能数据，进行数据面的性能评估，协助整体硬件配置规划。
 - iBMA和iBMC采用统一的数据模型和设备模型，设备位置和ID统一标识。



BOB是一个帮助带内网络远端主机通过带内组网与本机iBMC进行通信的转发工具。包括服务连接，如SSH、IPMI、HTTPS等；事件上报，如SNMP Trap等。

数据流有两种方式：

- (1) iBMA-->vNIC-->iBMC-->3rd
- (2) iBMA-->vNIC-->iBMC-->vNIC-->BoB-->3rd

iBMA功能介绍

- iBMA支持服务器CPU、内存、硬盘、网卡、IB卡、RAID卡、光模块等各个组件的状态监控、性能监控、事件监控。用户在系统上安装iBMA软件后，iBMA提供的主要功能如下表：

功能	介绍
系统信息获取	将操作系统、网卡、IB卡、硬盘、RAID卡等信息提供给iBMC。
文件上传	支持eSight和其它工具通过iBMC的Redfish接口将文件上传至iBMA运行的服务器。
驱动升级	支持eSight和其他工具通过iBMC的Redfish接口进行驱动升级。
系统状态监控	定时监控系统信息和健康状态的变化情况，用于上报给iBMC进行告警。
软件管理	支持FusionDirector和其他工具通过iBMC的Redfish接口进行软件的查询、安装、升级操作。
BOB	BOB支持服务管理（启动/停止/重启），不支持独立部署和工作。BOB提供了一个命令行bobCmd，用于启动/停止/重启BOB服务以及其他相关配置，转发配置支持在线配置，无需重启服务。配置支持持久化，即重启系统后不会丢失。
iBMC事件转储	Linux/Windows系统上支持将iBMC事件转储至系统日志文件中。
Redfish服务兼容HTTPS协议	Redfish服务支持以HTTPS或HTTP协议启动。
网卡配置追溯	在Linux系统下，iBMA会定时收集网卡的ethtool最新配置信息，文件存储在iBMA安装目录的“log/ethtool_log”目录下，可以使用SmartKit在收集操作系统日志时一并收集。
网卡故障检测	Linux系统支持Mellanox网卡的internal error检测，并记录在iBMA安装目录的“log/common.log”日志中。VMware系统支持检测网口状态短时间变化事件并上报更新给iBMC。
存储诊断	定时监控硬盘IO性能降低、硬盘Sense Code错误和文件系统只读的存储诊断信息并上报给iBMC。支持收集RAID卡下硬盘的日志，用于硬盘耗损分析。支持MD软RAID下硬盘的Firmware故障检测。
内存隔离	支持对操作系统出现故障的物理内存页，通过Redfish接口进行软隔离操作。

- iBMA同时提供了黑匣子功能，开启黑匣子功能后，能记录系统发生异常宕机时的状态信息和调试信息，可以辅助异常定位。

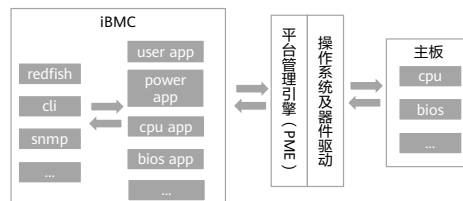
目录

1. 计算运维软件体系
- 2. 单机管理**
 - 带内管理
 - 带外管理
3. 集中管理
4. 近端运维
5. 远端运维

- 从技术的角度，网络管理可分为带外管理（out-of-band）和带内管理（in-band）两种管理模式。所谓带内管理，是指网络的管理控制信息与用户网络的承载业务信息通过同一个逻辑信道传送；所谓带外管理，是指网络的管理控制信息与用户网络的承载业务信息在不同的逻辑信道传送。

iBMC概述

- iBMC: Intelligent Baseboard Manager Controller, 智能基板管理控制器。
- 主要功能: 管理服务器, 如TaiShan 200、Altas 800等。
 - **对内**管理服务器上几乎所有的硬件, 主要包括: CPU、内存、Raid卡、硬盘、各类PCIE设备等, 管理项主要包括: 硬件信息、温度、电压、告警检测、硬件控制等。
 - **对外**提供管理接口, 主要包括: 命令行、WEB、IPMI、Redfish等。
- 运行环境: 独立于业务系统的嵌入式系统上, 硬件上使用ARM架构的自研芯片Hi1710/Hi1711; 软件栈包括RTOS及驱动、应用程序框架、应用程序等。

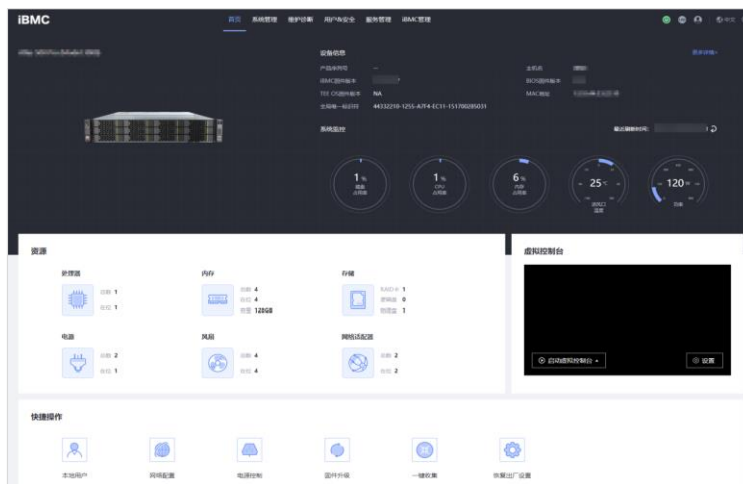


iBMC功能介绍

功能	介绍
丰富的管理接口	提供以下标准接口，满足多种方式的系统集成需求。 - DCMI 1.5接口。 - IPMI 1.5/IPMI 2.0接口。 - 命令行接口。 - Redfish接口。 - 超文本传输安全协议（HTTPS, Hypertext Transfer Protocol Secure）。 - 简单网络管理协议（SNMP, Simple Network Management Protocol）。
故障监控与诊断	可提前发现并解决问题，保障设备7*24小时高可靠运行。 - 系统崩溃时临终截屏与录像功能，使得分析系统崩溃原因不再无处下手。 - 屏幕快照和屏幕录像，让定时巡检、操作过程记录及审计变得简单轻松。 - 支持Syslog报文、Trap报文、电子邮件上报告警，方便上层网管收集服务器故障信息。
安全管理手段	- 通过软件镜像备份，提高系统的安全性，即使当前运行的软件完全崩溃，也可以从备份镜像启动。 - 多样化的用户安全控制接口，保证用户登录安全性。 - 支持多种证书的导入替换，保证数据传输的安全性。
系统维护接口	- 支持虚拟KVM（Keyboard, Video, and Mouse）和虚拟媒体功能，提供方便的远程维护手段。 - 支持RAID的带外监控和配置，提升了RAID配置效率和管理能力。 - 通过Smart Provisioning实现了免光盘安装操作系统、配置RAID以及升级等功能，为用户提供更便捷的操作接口。
多样化的网络协议	- 支持NTP，提升设备时间配置能力，用于同步网络时间。 - 支持域管理和目录服务，简化服务器管理网络。
智能电源管理	- 功率封顶技术助您轻松提高部署密度。 - 动态节能技术助您有效降低运营费用。
许可证管理	通过管理许可证，可实现以授权方式使用高级版的特性。iBMC高级版较标准版提供更多的高级特性，例如： - 通过Redfish实现OS部署。

iBMC功能：首页

- “首页”界面提供了：
 - 服务器的基本信息。
 - 虚拟控制台。
 - 服务器关键部件的信息及其快捷入口。
 - 系统监控项信息及其快捷入口。
 - 其他常用操作的快捷入口。



iBMC功能：系统管理

- 通过“系统信息”界面的功能，用户可以获取服务器的基本信息，包括产品信息、处理器、内存、网络适配器、传感器和其他部件的信息。



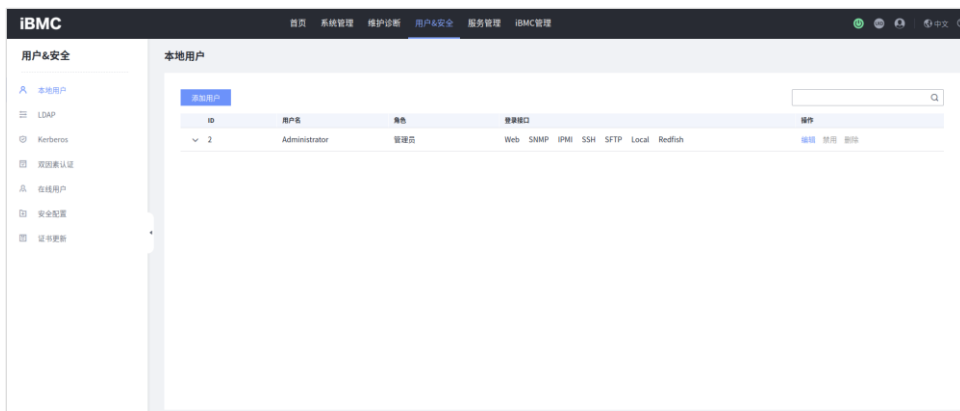
iBMC功能：维护诊断

- 通过“维护诊断”界面的功能，用户可以对服务器进行日常运维和故障处理，包含告警、录像截屏、系统日志（PCIe日志、黑匣子、系统串口数据记录）、iBMC日志（操作日志、运行日志、安全日志）等。



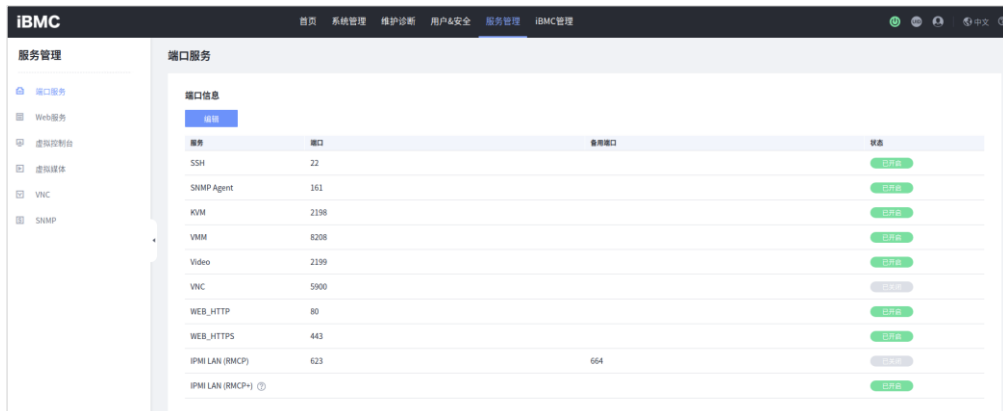
iBMC功能：用户&安全

- 通过“用户&安全”界面的功能，用户可以对服务器进行用户认证相关安全的管理，包含本地用户管理、LDAP、Kerberos、双因素认证、在线用户管理、用户登录安全配置、SSL证书管理等。



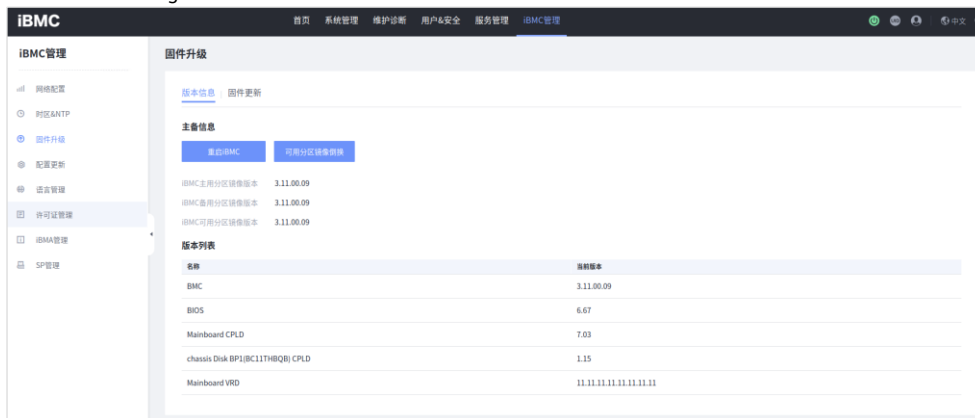
iBMC功能：服务管理

- 通过“服务管理”界面的功能，用户可以对服务器的对外的各种服务进行管理，包含服务端口管理、Web服务管理、虚拟控制台、虚拟媒体、VNC、SNMP等。



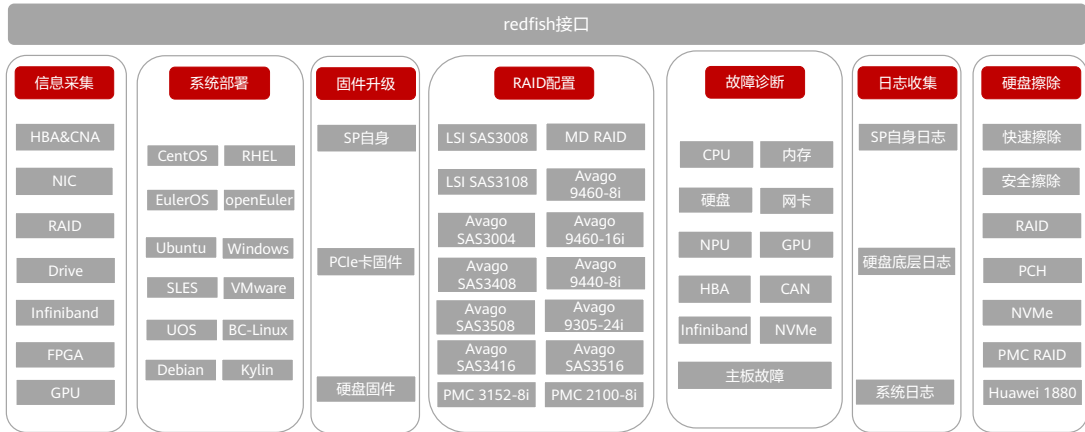
iBMC功能：iBMC管理

- 通过“iBMC管理”界面的功能，用户可以对服务器的iBMC进行升级、配置更新、网络等相关配置，同时也可对iBMA和Smart Provisioning进行管理。



Smart Provisioning功能全景图

Smart Provisioning作为iBMC能力向带内的延伸，通过运行在业务内存中的小系统，向用户提供信息采集、系统部署、固件升级、RAID配置和故障诊断等功能，同时这些能力通过iBMC向用户开放redfish接口。



Smart Provisioning功能界面



- ✓ 安装操作系统
- ✓ 配置RAID
- ✓ 升级



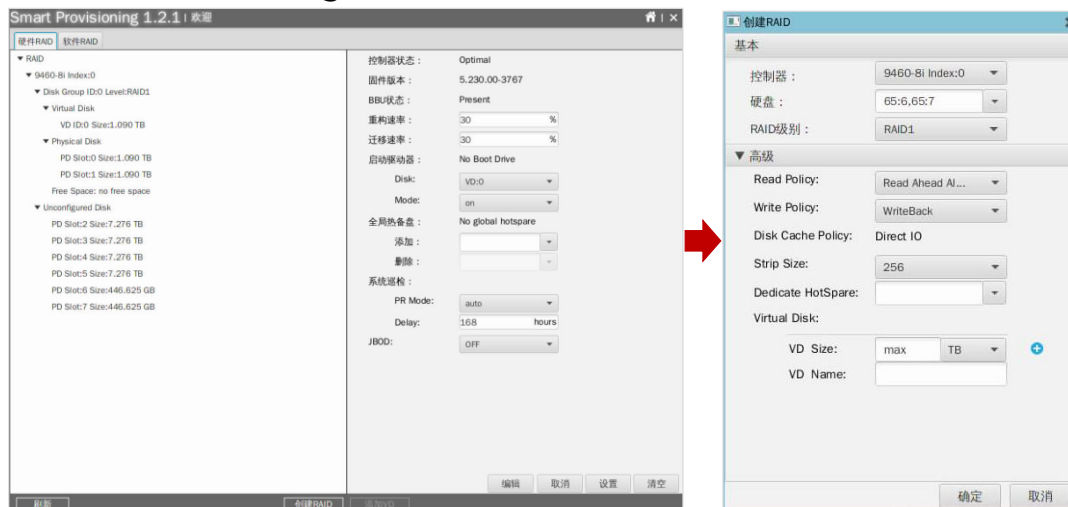
- ✓ 硬件诊断
- ✓ 日志收集
- ✓ 硬盘擦除

Smart Provisioning功能介绍：安装操作系统



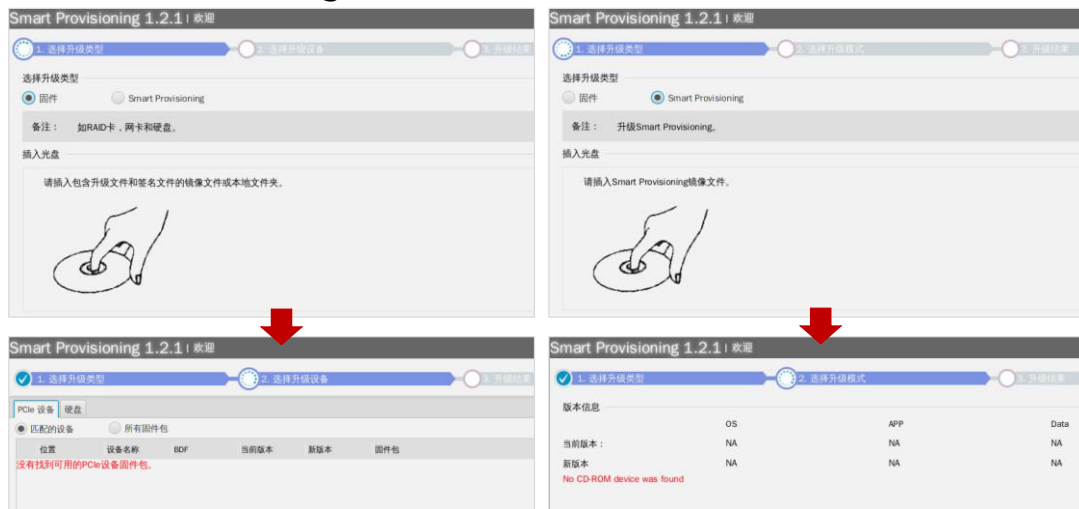
- 支持使用UEFI模式安装操作系统。
- 支持安装CentOS、EulerOS、Ubuntu、SUSE Linux Enterprise Server（SLES）、openEuler、BC-Linux、Kylin、Debian和NeoKylin等主流操作系统。

Smart Provisioning功能介绍：配置RAID



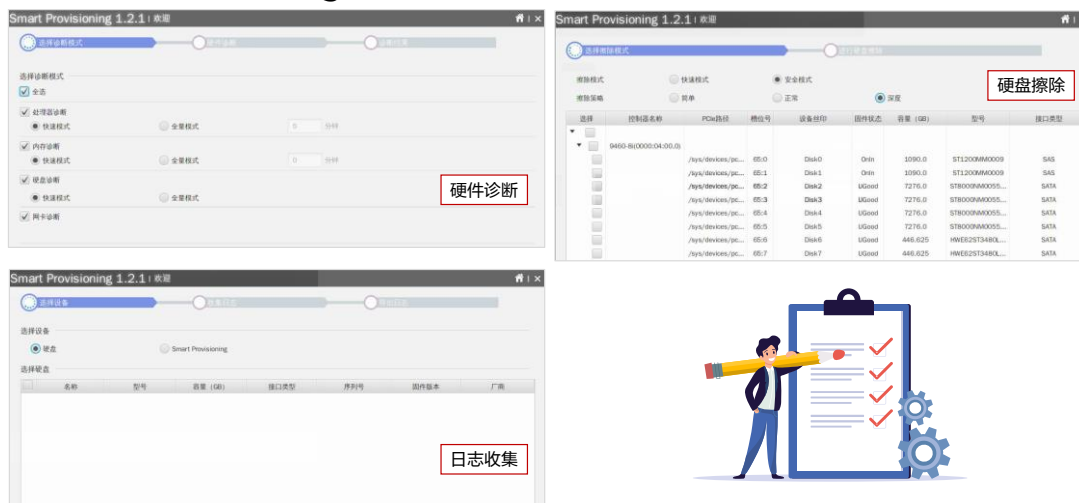
- 支持配置Avago SAS3408、Avago SAS3508、Avago SAS3416iMR、Avago MegaRaid SAS 9460-16i/9460-8i/9440-8i/9560-8i RAID卡。
- 支持配置软件RAID。

Smart Provisioning功能介绍：升级



- 支持PCie设备包括RAID卡和网卡的固件升级。
- 支持硬盘（SAS，SATA和NVMe）固件升级。
- 支持Smart Provisioning固件升级。

Smart Provisioning功能介绍：工具箱

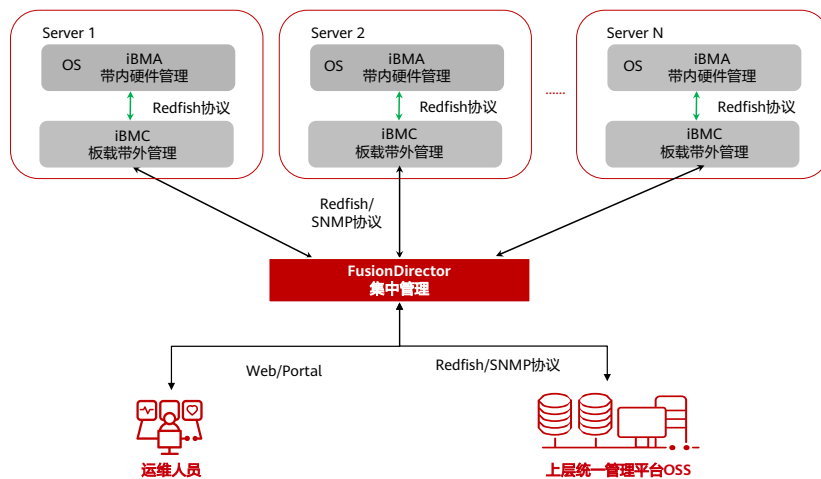


- 硬件诊断：
 - 支持对处理器、内存，硬盘和网卡进行诊断。
- 日志收集：
 - 支持收集硬盘的日志。
 - 支持Toshiba和HGST的硬盘。
 - 支持收集Smart Provisioning日志。
- 硬盘擦除：
 - 支持对RAID卡，软件RAID和南桥直通（AHCI）下硬盘和NVMe硬盘的擦除。
 - 快速模式对逻辑盘的分区进行擦除。
 - 安全模式对物理盘进行全盘擦除。

目录

1. 计算运维软件体系
2. 单机管理
- 3. 集中管理**
 - FusionDirector
 - eSight
4. 近端运维
5. 远端运维

FusionDirector概述



- FusionDirector作为计算设备集中管理的运维大脑，解决服务器批量集中化运维问题，降低运维OPEX，提升数据中心服务器整体运行稳定性。

智能自动部署

智能资产管理

智能版本管理

智能故障预测

智能功耗管理

服务器设备纳管

批量导入方式：适用于批量设备已经配置好IP地址，且设备的帐号、IP地址等信息都已经知道的情况下精确纳管。

批量导入

1. 登录FusionDirector WebUI。
2. 进入设备批量导入页面下载设备导入模板。
3. 填写设备导入信息。
4. 在设备批量导入页面导入填写好的设备导入模板。

手动添加方式：适用于批量设备已经配置好IP地址，且设备的帐号相同、IP地址分布比较连续的情况下的模糊纳管。

手动添加

1. 登录FusionDirector WebUI。
2. 进入设备手动添加页面。
3. 填写需要纳管的批量设备网段信息和帐号信息。
4. 在设备批量导入页面导入填写好的设备导入模板。

DHCP自动分配IP方式：适用于批量设备处于DHCP分配模式，且设备的帐号、SN或MAC地址已经知道的情况下纳管。

限制：需要在网络中搭建DHCP中继服务器，且规划好每个设备要分配的IP地址。

DHCP自动分配IP

1. 登录FusionDirector WebUI。
2. 进入配置规划页面下载配置规划模板。
3. 填写需要纳管的批量设备的SN、IP地址、帐号信息。
4. 在配置规划页面导入填写好的配置规划模板。
5. 打开DHCP服务开关。

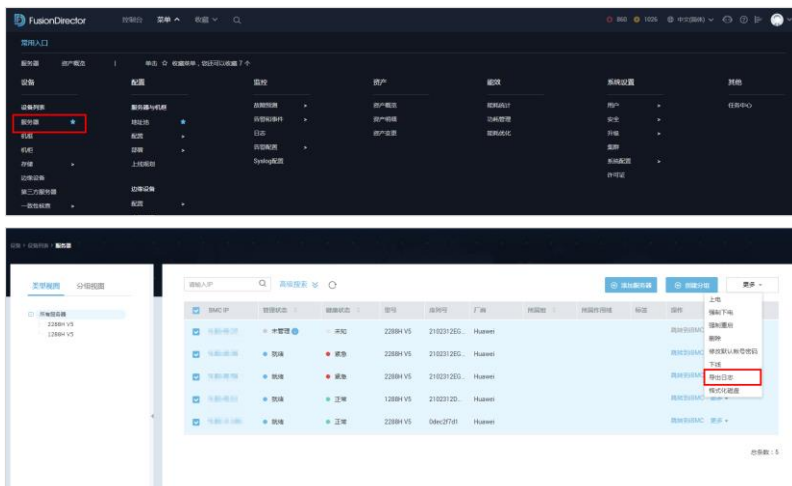
SSDP自动发现方式：IP-SN模式适用于已知设备SN、帐号信息，将批量设备配置为导入数据中的IP；IP池模式适用于未知设备信息情况下对设备进行IP地址随机分配的纳管。

限制：FusionDirector必须和设备处于同一个二层网络，不能跨路由。

SSDP自动发现

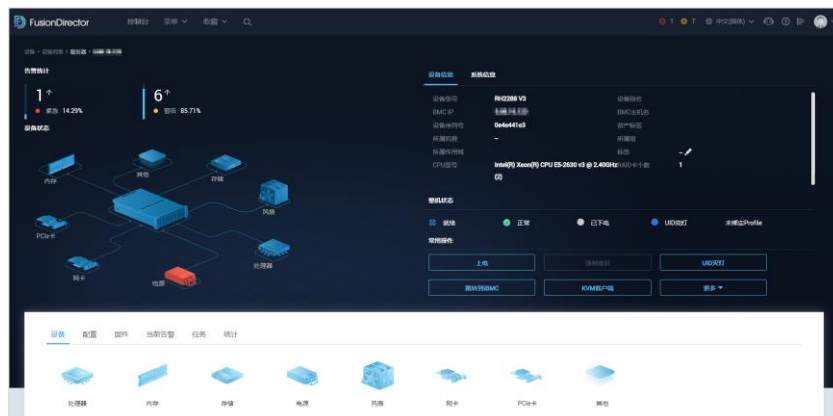
1. 登录FusionDirector WebUI。
2. 进入设备自动发现页面。
3. 选择IP-SN模式下模板导入需要纳管的设备信息或选择IP池模式配置好要分配的IP段。
4. 开启SSDP服务。

服务器日志收集



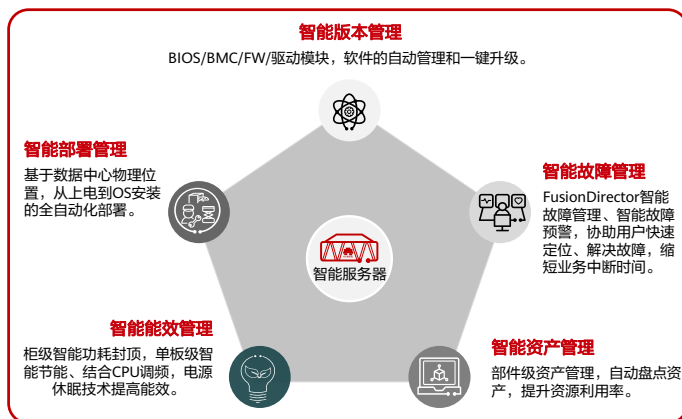
FusionDirector支持对接入的服务器进行信息展示和管理。在服务器管理界面支持一键导出日志，一次最多支持导出100台服务器日志，方便用户操作。

服务器信息展示和操作



- 1 支持展示的资源主要包括服务器的处理器、内存、存储、电源、风扇、网卡、固件、告警等信息。
- 2 支持的操作主要有服务器上电、强制重启、UID点灯/灭灯、跳转到iBMC、KVM客户端、删除、刷新、修改默认帐号密码、BMC IP配置、格式化磁盘等操作。

五大智能特性

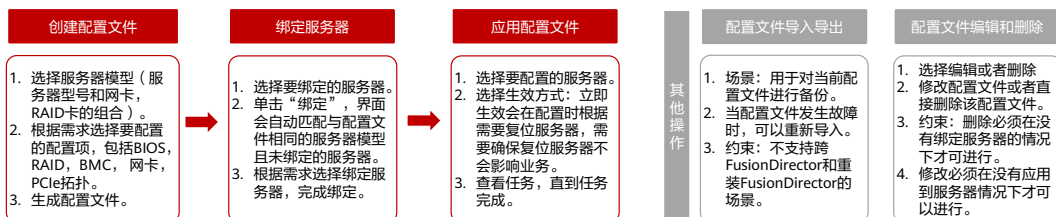


智能部署：服务器配置文件



服务器型号1	+	RAID卡型号1	+	网卡型号1	=	设备模型1
服务器型号2	+	RAID卡型号2			=	设备模型2
服务器型号3	+	RAID卡型号3	+	网卡型号2	=	设备模型3

- ① 一个配置文件可以同时绑定到多台服务器模型相同的服务器上。
- ② 不同服务器模型的服务器，需要创建不同的配置文件，才可配置。
- ③ 绑定和应用是分离的，必须先绑定才可应用到服务器。



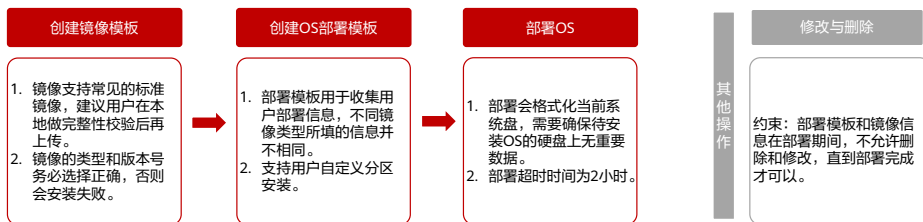
智能部署：OS部署

场景：适合于开局、运维过程中的带内OS安装，支持主流操作系统。

准备工作：

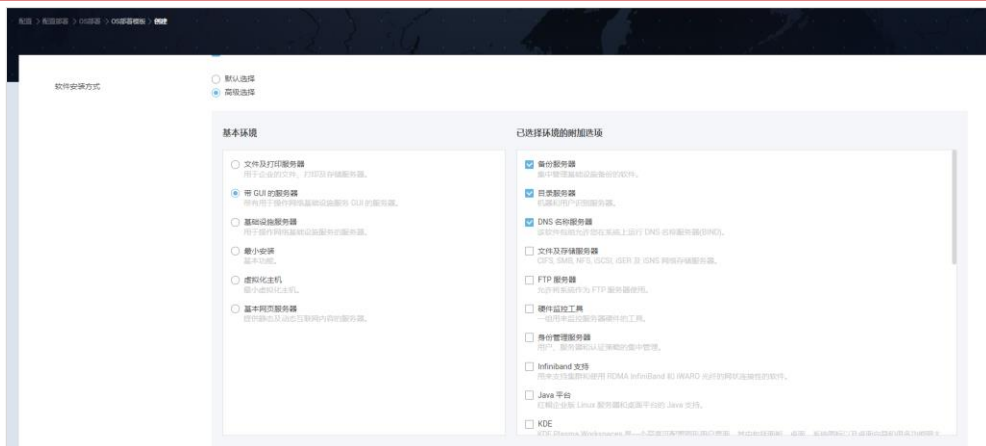
- 1 当前仅支持含有Smart Provisioning的服务器。
- 2 该部署属于带外部署，不同于PXE，故不需要DHCP及TFTP的参与，只要被纳管，则网络部署就已经符合条件。
- 3 部署前请务必配置好RAID及启动盘，否则会导致部署失败，可使用服务器配置文件中的配置选项来完成RAID配置。
- 4 请检查当前待部署服务器的RAID卡是否在用户指南中支持。
- 5 OS部署属于高级特性，请确保在待部署设备上安装了License或者FusionDirector已经安装了支持OS部署的License。

规格：当前最多支持同时部署30台设备。



智能部署：选择软件列表

在OS部署模板创建界面，可以对OS部署时一起安装的软件列表进行定制化选择，支持的操作系统包含：KylinV10.*、openEuler*、EulerOS*、RHEL7U*。



智能部署：自动上线规划



准备工作

1. 创建服务器模型。
2. 创建服务器配置文件。
3. 创建OS部署模板。

创建基于位置的规划

按位置规划依赖智能机柜iRM来获取服务器部署在机柜的U位位置。

1. 设置规划模板基础信息。
2. 规划服务器位置。
3. 规划服务器配置。

创建基于SN/MAC的规划

按SN/MAC规划依赖DHCP来获取服务器上架消息，需要待上架设备跟FusionDirector在一个DHCP广播域内。

1. 设置规划模板基础信息
2. 导入规划服务器信息
3. 规划服务器配置

上线规划开启与关闭

其他操作

只有当上线规划开启后才会激活规划中待上架的数据，关闭开关后不会影响已经激活的设备上线。当开启位置规划时，基于SN/MAC规划的激活开关会自动关闭。当开启基于SN/MAC规划时，基于位置规划的激活开关会自动关闭。

智能资产管理

当前资产管理痛点

人工错误多

盘点耗时长

10%以上
的人工操作错误

提前2个月进行盘
点审核，单次耗
费3人/月以上

丢失资产

资产定位难

0.25%以上
的资产丢失

核对一个错误资产
需近2小时

如何破解困局？

华为解决方案

柜级资产智能管理，硬件实现U位资产自动识别主动上报，入库、盘点免人工干预。



效果收益



自动

Automation

- 免人工干预。
- 入库盘点效率提升95%以上。



精准

Accurate

- 部件级资产管理（CPU、内存、硬盘）。
- 100%的资产准确率。

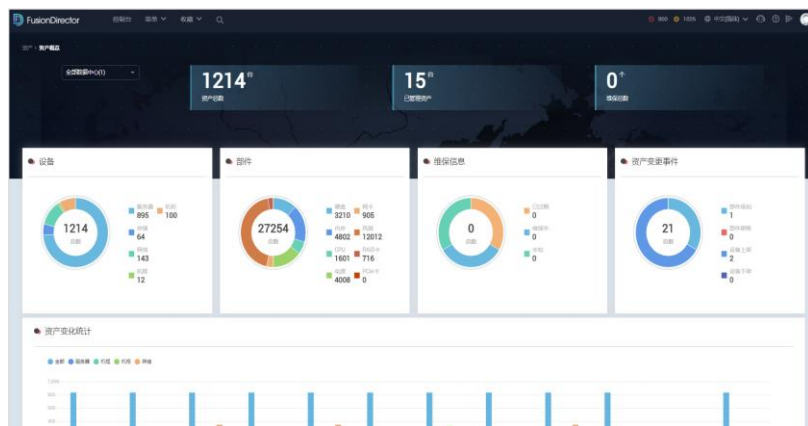


安全

Security

- 全面资产信息报表。
- 实时追踪资产变更情况，防止资产损失。

智能资产管理：资产概览



- 1 支持资产统计总量、服务器、存储、网络、机框、机柜等设备的总数统计。
- 2 支持资产变更概览，统计设备的上架、下架和部件更换情况。
- 3 支持资产上架使用年限，并按照服务器、存储和网络设备展示。
- 4 支持按照数据中心展示资产总数和空间利用率。

智能资产管理：资产变更

- 在“资产变更统计”界面，展示全部或指定数据中心的设备和部件一年内的变更统计情况。在页面下方的变更记录表中，体现变更的设备或部件类型、变更时间、变更事件类型、序列号、型号、所在位置、所在数据中心名称、资产编号以及变更事件的描述信息。

资产类型	变更时间	类型	序列号	型号	资产位置	数据中心	资产编号	描述
内存	+0000 +..	部件添加	2102311TYANDJ9000020	2288H V5	LocalDataCenter-X1-Y1-U5	LocalDataCenter		Model 2288H V5-sp-add
硬盘	+0000 +..	部件添加	2102311TYANDJ9000020	2288H V5	LocalDataCenter-X1-Y1-U5	LocalDataCenter		Model 2288H V5-sp-add
CPU	+0000 +..	部件添加	2102311TYANDJ9000020	2288H V5	LocalDataCenter-X1-Y1-U5	LocalDataCenter		Model 2288H V5-sp-add
电源	+0000 +..	部件添加	2102311TYANDJ9000020	2288H V5	LocalDataCenter-X1-Y1-U5	LocalDataCenter		Model 2288H V5-sp-add
网卡	+0000 +..	部件添加	2102311TYANDJ9000020	2288H V5	LocalDataCenter-X1-Y1-U5	LocalDataCenter		Model 2288H V5-sp-add
风扇	+0000 +..	部件添加	2102311TYANDJ9000020	2288H V5	LocalDataCenter-X1-Y1-U5	LocalDataCenter		Model 2288H V5-sp-add
RAID卡	+0000 +..	部件添加	2102311TYANDJ9000020	2288H V5	LocalDataCenter-X1-Y1-U5	LocalDataCenter		Model 2288H V5-sp-add

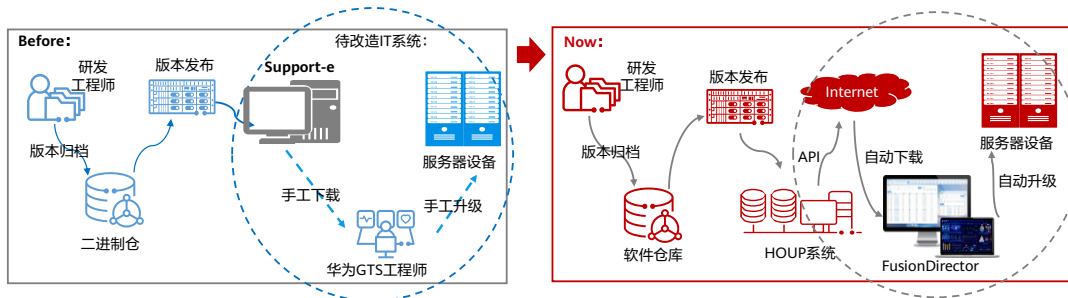
总页数: 7

智能资产管理：资产明细

- 在“资产明细”界面，展示全部或指定数据中心的资产统计信息，包括设备和部件，可以根据需要选择查看设备详细信息或部件详细信息。

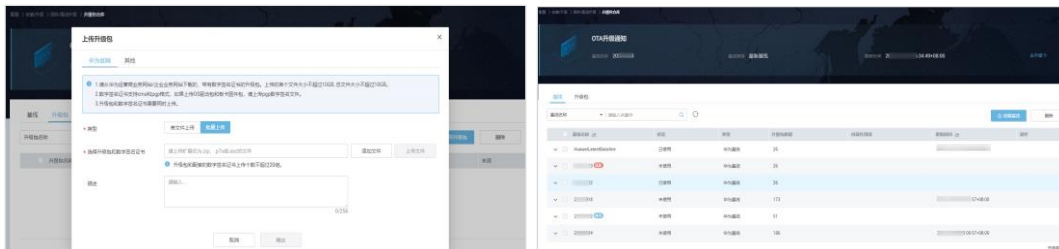
智能版本管理

- 自动下载，自动检测不同型号服务器的固件、驱动及OS版本，全面化管理固件版本。
- 无需运维工程师手动下载固件、手动上传固件、手动升级固件。
- 独立带外通道升级，不占用业务带宽。
- OTA：支持新基线或推荐基线的弹窗通知/基线修改说明功能/预警的推送和展示/支持升级提前通知/升级报告导出。



智能版本管理：升级包仓库

- 升级包仓库：通过“配置>安装/升级>固件/驱动升级>升级包仓库”可以查询本地升级包列表以及基线列表，用于管理升级包和基线信息。



支持方式：

- 从本地导入的升级包，支持单文件上传、批量上传。
- 华为基线：从HOUUP上自动下载的基线，基线里的升级包已经符合版本配套关系以及兼容性关系。
- 自定义基线：由用户手动创建的基线，可自由配置。例如，当华为基线里的升级包不满足升级需求时，可使用自定义基线，将华为基线与本地导入的升级包进行组合创建一个自定义基线。

智能版本管理：升级计划

- 升级计划：通过“配置>安装/升级>固件/驱动升级>升级计划”可以查询创建的升级计划列表。用户可以创建、修改、启动、停用升级计划。

升级计划

如果没有操作权限，请联系系统管理员检查是否对升级计划中的所有设备都具有操作权限。

创建计划 删除

名称	状态	类型	升级时间	创建时间	更新时间	操作
2	138 OTA 已完成	OTA升级计划	立即执行	20:00:00		修改 应用 查看任务 导出报告
test	未应用	普通升级计划	每周 星期一 00:00:00 执行	20:00:00	20:00:00	修改 应用 查看任务 导出报告

总条数: 2

升级计划类型：

- 普通升级计划：用户根据升级需要，手动创建的计划。
- OTA升级计划：用户接收HUP推送的新基线或推荐基线创建的升级计划。
- 周期升级计划：对新设备或备件更换的设备进行周期检查的升级计划。

智能版本管理：设备版本状态

- 设备版本状态：通过“配置>安装/升级>固件/驱动升级>设备版本状态”可以查询设备版本状态信息，并执行手动生效。

系统 > 安装/升级 > 固件/驱动升级 > 设备版本状态

设备版本状态 0/0000的固件升级不支持使用生效按钮生效，用户需自行生效

地址 请输入关键字 🔍 🔄 🔊 生效

☐ 地址	类型	型号	健康状态	序列号	所属组	操作
^ ☐ 未激活设备	Server	Z28BH V5	警告	1d5d78209	-	生效

组件	型号	厂商	当前版本	目标版本
ActiveBMC	iBMC	Huawei	5.06	
BackupBMC	iBMC	Huawei	5.06	
Bios	BIOS(Z28BH V5 10GE SFP+)	Huawei	7.55	7.79 ↑
Driver	x86_64	CentOS	7.6	
MainBoardCPLD	CPLD(MainBoard)	Huawei	2.10	
PS1	pac900s12-be	huawei	dc:128 pfc:126	
PS2	pac900s12-be	huawei	dc:128 pfc:126	
SmartProvisioning	SmartProvisioning	huawei	1.25	
chassisDiskBP1CPLD	CPLD(12*3.5 SAS/SATA, Expander)	Huawei	1.11	

智能版本管理：软件部署

- 支持软件的安装、升级、回退和卸载：通过“配置>安装/升级>软件部署”进行软件包仓库管理、计划管理以及设备软件版本信息的查询，软件部署的流程与固件/驱动升级流程类似。

The screenshot displays the '配置 > 安装/升级 > 软件部署 > 软件包仓库' (Configuration > Installation/Upgrade > Software Deployment > Software Package Repository) page. On the left, there is a '基本信息' (Basic Information) form with fields for '名称' (Name), '描述' (Description), '类型' (Type), and '应用' (Application). Below this is the '目标设备' (Target Devices) section with a search bar and a list of devices. The main area shows the '软件包仓库' (Software Package Repository) with a search bar and a table of device software versions.

名称	名称	管理状态	健康状态	型号	序列号	设备厂商	操作	
☐	192.168.67.173	Computer System	部署	Z288H V5	2102312EGJ15.8000006	Huawei	生效 回退	
☒	192.168.19.72	3e9d1082-023e-4f5a-b...	未管理	未知	Atlas 200	3e9d1082-023e-4f5a-bc99-607da9e49472	huawei	生效 回退
☒	192.168.19.73	154a32f5-9d5f-44a3-8...	未管理	未知	Atlas 200	154a32f5-9d5f-44a3-810c-41b5f2284629	huawei	生效 回退
☐	192.168.53.10	test_zij-3-nodeid	部署	正常	Atlas500	asdf45d3d3dgnest_zj3	Huawei	生效 回退

软件名称	当前版本	待生效版本	可回退版本	软件厂商
HA	20.1.0	20.2.0		Huawei Technologies Co., Ltd

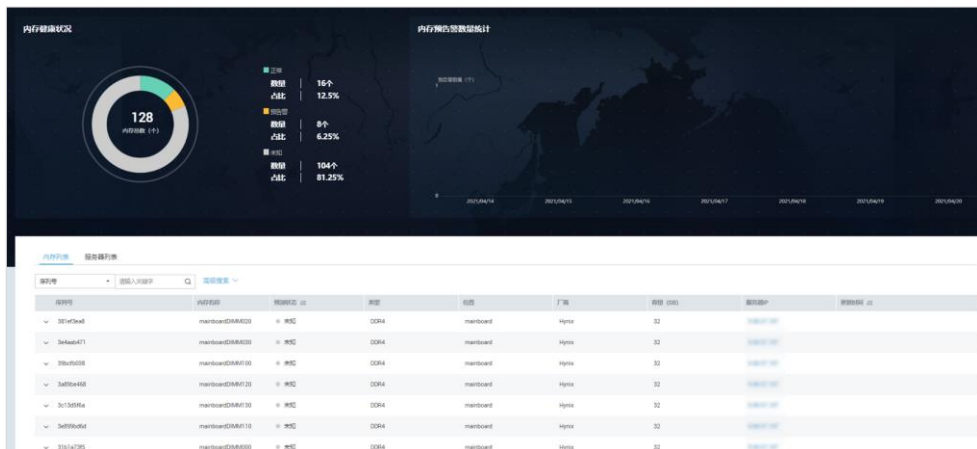
智能故障管理：硬盘故障预测

- 硬盘故障预测：展示对已纳入系统管控的所有硬盘的健康状况、硬盘风险预测图表，展示硬盘风险和硬盘统计的详细信息，帮助您更好地了解硬盘状态。



智能故障管理：内存故障预测

- 内存故障预测：支持查看内存的健康状态，以及自愈操作的历史记录。



智能能效管理

数据中心能耗痛点

租赁维护费用高：

机柜租赁
80+%

机房电费
78+%

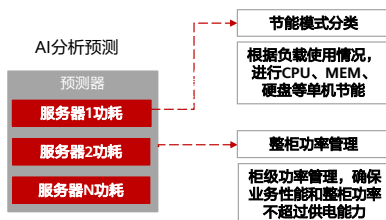
利用率低：

机柜空间
40%~50%

机柜供电
利用率
约50%

提升机柜利用率，
降低电费是当前的TOP工作

华为解决方案



AI分析预测整柜服务器运行功率，配合整柜功率封顶功能，业务模型不变情况下提升机柜部署密度，机柜利用率提升。

效果收益

机柜空间
利用率提升

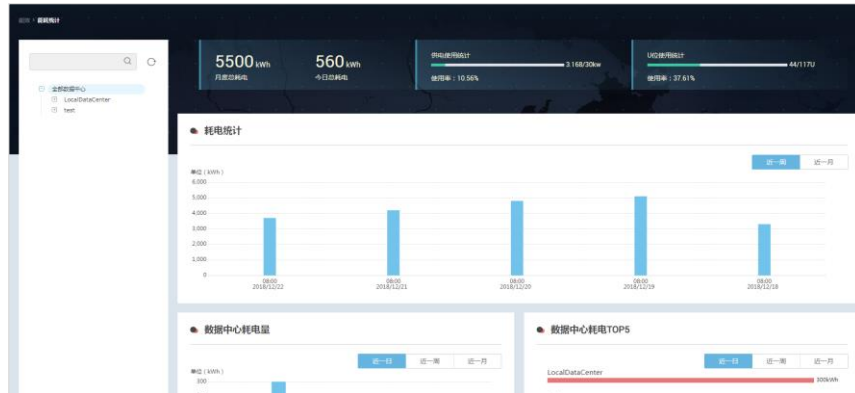
机柜供电
利用率提升

租赁费用下降

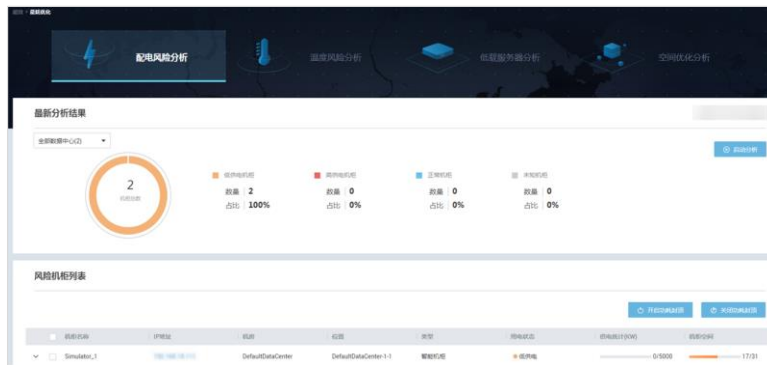
机房电费下降

智能能效管理：能耗统计

- FusionDirector对已管理的所有数据中心进行能耗统计，持续记录每台服务器的功耗数据。用户可以根据需要选择查看全部数据中心、单个数据中心、数据中心中的单个机柜或单台服务器的功耗数据，根据这些数据推理未来的功耗需求，调整该设备功耗封顶值。



智能能效管理：能耗优化



功能	描述
配电优化分析	统计出长期供电利用率比较高的机柜，并可以对机柜进行批量功耗封顶的操作。
温度风险分析	统计入风口温度过高的服务器信息，并可以对服务器进行下电操作。
负载优化分析	统计出长期低载服务器信息，并可以对低载服务器进行批量操作。
空间优化分析	统计长期空间利用率比较高的机柜。

智能能效管理：功耗管理

- 在“功耗管理”界面中，显示FusionDirector所管理的智能机柜的功耗封顶情况，并提供封顶功耗、封顶策略的设置接口。

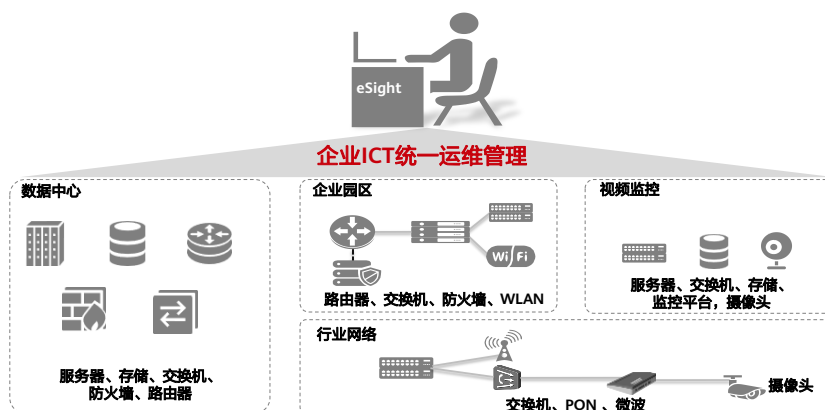
名称	位置	功耗 (W)	额定功率 (W)	功耗封顶 (W)	功耗封顶状态	功耗封顶策略	功耗封顶次数	功耗封顶时间 (min)	策略
Rack205_1	wuhanDC 21-1	900/1500	1500	600	已开启	功耗封顶	5	30	设置功耗封顶策略
Rack205_2	wuhanDC 21-1	900/1500	1500	600	已开启	功耗封顶	5	30	设置功耗封顶策略
Rack205_3	wuhanDC 21-1	900/1500	1500	600	不可配置	未知	5	30	设置功耗封顶策略
Rack205_4	wuhanDC 21-1	900/1500	1500	600	已开启	功耗封顶	5	30	设置功耗封顶策略
Rack205_5	wuhanDC 21-1	900/1500	1500	600	已开启	功耗封顶	5	30	设置功耗封顶策略
Rack205_6	wuhanDC 21-1	900/1500	1500	600	已开启	功耗封顶	5	30	设置功耗封顶策略
Rack205_7	wuhanDC 21-1	900/1500	1500	600	已开启	功耗封顶	5	30	设置功耗封顶策略

目录

1. 计算运维软件体系
2. 单机管理
- 3. 集中管理**
 - FusionDirector
 - eSight
4. 近端运维
5. 远端运维

eSight概述

可视化运维、极简运维、智能化运维、融合共管



- eSight是面向企业的一体化融合运维管理解决方案，可实现交换机、路由器、WLAN、防火墙、视频监控、服务器、存储、微波、PON设备以及服务器操作系统的统一管理，为企业ICT设备提供集中化管理、可视化监控、智能化分析等功能，有效帮助企业提高运维效率、降低运维成本、提升资源使用率，有效保障企业ICT系统稳定运行。标准版提供完整的服务、管理规模和基础技术指标，多种部署形态，满足不同的场景要求。

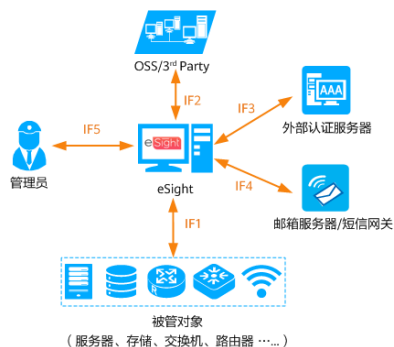
eSight架构



- 多厂商设备适配，轻松实现全网设备统一管理
 - 支持华为交换机、路由器、WLAN、防火墙、视频监控、服务器、存储、微波、PON设备以及服务器操作系统的统一管理，并预集成对HP、Cisco、H3C等第三方主流设备的管理能力。
- 可靠性高、管理容量大
 - 基于服务化架构，支持单机、双机主备、多节点多活架构部署，服务级故障隔离，降低业务中断时长。
 - Scale-out架构，支撑后续按需扩展。灵活支持从几十、几百、几千、几万不同管理规模，支持小规模、大规模不同场景一套共管诉求。
- 轻量级、Web化，减少系统维护与升级成本
 - eSight采用B/S架构，客户端无需安装任何插件，可随时随地访问。当系统升级或维护时只需更新服务器端软件，减少系统维护与升级成本。
- 开放、易集成
 - 开放式架构，基于RESTful API接口，方便外部系统对接集成。

- eSight采用服务化架构，逻辑架构包括开放集成服务、运维分析服务、设备管理服务、基础管理服务、南向设备接入服务。

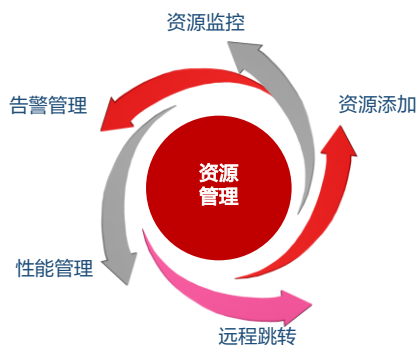
eSight外部接口



- IF1: 南向接口，用于eSight与设备之间的对接，完成eSight对设备的管理功能。
- IF2: 北向接口，eSight支持与上层管理系统的集成。上层OSS或第三方系统可通过RESTful、SNMP、FTP或Kafka接口，获取eSight系统中管理的资源、告警、性能等信息。eSight支持与Syslog服务器对接，将eSight日志转发到Syslog服务器。
- IF3: 外部认证服务器集成接口，eSight支持与企业自建认证服务器对接，使用认证服务器中的用户登录eSight系统，支持的对接协议包括LDAP、RADIUS和SSO。
- IF4: 消息通知接口，eSight支持与企业已有邮箱服务器或短信网关对接，将收到的重要告警、事件、报表等以电子邮件或短信方式通知客户。
- IF5: 用户接口，即界面UI，系统管理员及操作人员通过此接口使用eSight所提供的各项功能。

- eSight支持的南向接口类型，包括SNMP、STelnet、FTP（服务端）、SFTP（客户端/服务端）、IPMI、HTTP（客户端）/HTTPS（REST/Redfish客户端、服务端）、WebSocket、SocketPing、ICMP、WebPing、WebTrace、SSDP、syslog、WebTelnet、Netconf、TR069。

eSight服务器特性介绍



特性概述

- 资源添加：支持华为自研服务器和第三方服务器管理。
- 资源监控：支持服务器基本信息（包括型号、状态、产品序列号、功率等），系统信息（包括iBMA信息、OS信息、域名等），部件信息（包括主板、CPU、GPU、内存、硬盘、电源、风扇、网卡、PCIe卡、Raid卡、MEZZ卡、网口、传感器等）进行状态监控。
- 告警监控：支持华为自研服务器和第三方服务器告警监控。
- 性能监控：对服务器CPU利用率、内存利用率、硬盘利用率等指标进行实时性能和历史性能监控。
- 远程跳转：跳转到服务器的iBMC管理界面或远程控制台。

- 资源包括：服务器设备资源、部件、告警、性能等；
- 设备包括：机架服务器、刀片服务器、高密度服务器、异构计算服务器、KunLun服务器、第三方服务器等；

资源添加

- 支持通过SNMP和Redfish协议接入服务器，提供单个添加、批量导入和自动发现方式添加服务器。

设备类型	接入方式	发现协议	管理协议
机架服务器	单个添加/批量导入/自动发现/SSDP发现	Redfish/SNMP	Redfish/SNMP
刀片服务器	单个添加/批量导入/自动发现/SSDP发现	Redfish/SNMP	Redfish/SNMP/SSH
高密服务器	单个添加/批量导入/自动发现/SSDP发现	Redfish/SNMP	Redfish/SNMP
TaiShan服务器	单个添加/批量导入/自动发现/SSDP发现	Redfish/SNMP	Redfish/SNMP
异构服务器	单个添加/批量导入/自动发现/SSDP发现	Redfish	Redfish

The left screenshot shows the 'Basic Information' tab. It includes fields for 'IP地址' (IP Address), '名称' (Name), and '接入协议' (Access Protocol). Under '接入协议', there are radio buttons for 'Redfish' and 'SNMP'. Below this, there are dropdown menus for 'SNMP协议版本' (SNMP Protocol Version) and 'SNMP社区字符串' (SNMP Community String). There are also fields for '用户名' (Username), '密码' (Password), and '超时时间 (秒)' (Timeout (s)).

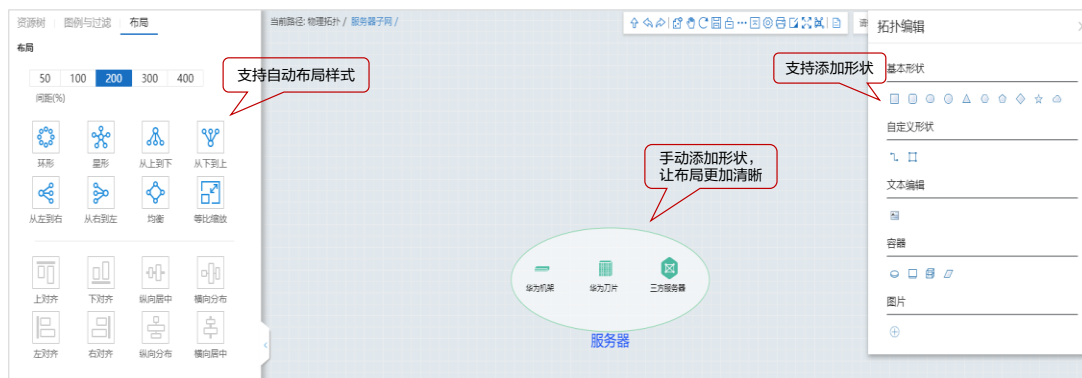
The right screenshot shows the 'Protocol Settings' tab. It includes fields for '协议版本' (Protocol Version), 'SNMP协议' (SNMP Protocol), 'Redfish协议' (Redfish Protocol), and 'REST协议' (REST Protocol). There are radio buttons for '选择协议类型' (Select Protocol Type) and '选择发现模式' (Select Discovery Mode). Below this, there are fields for '认证方式' (Authentication Method), '用户名' (Username), '密码' (Password), and '超时时间 (秒)' (Timeout (s)).

57 Huawei Confidential



- 单个添加
 - 一次添加一个设备。适用于网络正常运行过程中，需要增加一台或几台少量设备的情况。
- 批量导入
 - 通过Excel表一次添加不超过500台设备。适用于需要添加多个网段中零散分布的设备，或者添加同一网段中的批量设备。
- 自动发现
 - 通过指定网段的方式，快速添加网段内批量设备。
- SSDP发现
 - 当服务器设备未配置iBMC管理IP地址等网络配置信息时，可以通过SSDP（Simple Service Discovery Protocol）的发现方式，根据服务器设备的SN、接入协议等信息发现设备和配置管理IP地址等网络信息，并接入eSight进行管理。

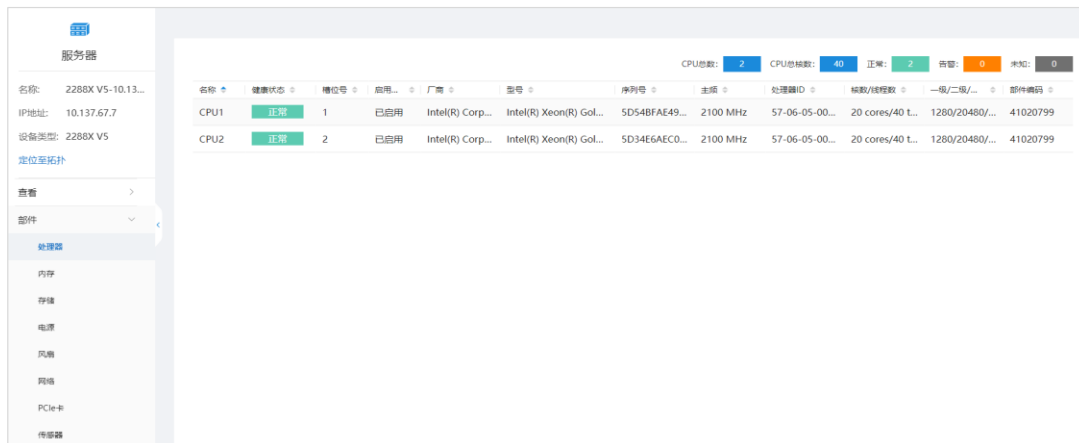
资源监控：拓扑布局





- 概览设备资源、查看设备基本信息（名称、型号、IP地址、序列号、固件版本、BIOS版本、电源状态、健康状态、内存容量、CPU型号）、设备面板、当前告警和历史告警。

资源监控：部件信息



- 部件详情展示信息包括如下信息：

- 处理器信息：名称、健康状态、启用状态、厂商、型号、主频、处理器ID、核数/线程数、一级/二级/三级缓存、部件编码、CPU总数、CPU总核数。
- 内存信息：名称、健康状态、位置、厂商、容量、主频、序列号、类型、最小电压、RANK（列）、位宽、技术、部件编码、内存总数、内存总容量。
- 电源信息：名称、健康状态、厂商、型号、固件版本、协议、额定功率、输入模式、冗余模式、工作模式。
- 风扇信息：名称、健康状态、部件编码、控制模式。
- 网络信息：网卡名称、健康状态、网卡位置、网卡类型、厂商、型号、芯片厂商、芯片型号、固件版本、驱动名称、驱动版本。
- PCIe卡：名称、健康状态、槽位号、类型、厂商、描述、固件版本。
- 传感器：名称、状态、单位、紧急上门限、紧急下门限、严重上门限、严重下门限、轻微上门限、轻微下门限。

告警管理

告警管理界面展示了服务器的告警信息，包括告警列表、告警详情和告警统计。

告警列表：

操作	级别	名称	告警源	首次发生时间	定位信息
☐	严重	CCU线或下在位	210231V0000000000009	23:1...	事件主体=cable
☐	严重	电源输入类别 电源通讯异常	H3	10:2...	
☐	严重	电源输入丢失	server2345	15:3...	事件主体=PSU2
☐	严重	性能阈值告警	test3	10:1...	指标名称: 已使用百分比, 分区名称: /dev/mapper/vg_root-var_log, 挂载
☐	严重	网管服务器与网元通讯异常	57703	10:0...	
☐	严重	网管服务器与网元通讯异常	252.78	05:3...	
☐	严重	第三方智能视觉平台断连	OSS	09:4...	域名=domainName0, 第三方平台名称=platformName4
☐	严重	第三方智能视觉平台断连	OSS	09:4...	域名=domainName0, 第三方平台名称=platformName3
☐	严重	第三方智能视觉平台断连	OSS	09:4...	域名=domainName0, 第三方平台名称=platformName2
☐	严重	SNMP连接失败	FS_az1.dc1.domaintsuki.com	09:4...	restype=vim,地址=10.136.154.179,协议=SNMP,端口=20161
☐	严重	组件故障	FS_az1.dc1.domaintsuki.com	09:5...	restype=host,az=az1.dc1,id=host=FS002,主机=host01,云服务=OpenStack
☐	严重	HAProxy代理服务不可用	FS_az1.dc1.domaintsuki.com	09:3...	restype=vim,服务故障=HAProxy Detects Abnormal Heartbeats of Ba
☐	严重	OpenStack OM与SNMP管理站连...	FS_az1.dc1.domaintsuki.com	09:2...	restype=vim,云服务=OpenStack OM,SNMP管理站ID=106

告警详情：

拓扑上设备告警统计

网元管理器监控告警统计

返回搜索列表

server2345

网元地址: 10.136.155.234 类型: 2288H VS

告警: 1 严重 0 次要 0 轻微

常用操作: 无相关操作

性能管理

- 实时性能展示指标信息包括：电源功率、设备CPU利用率、进风口温度、热量、功耗、硬盘利用率、风扇转速、设备内存利用率、网络带宽资源占用率等。

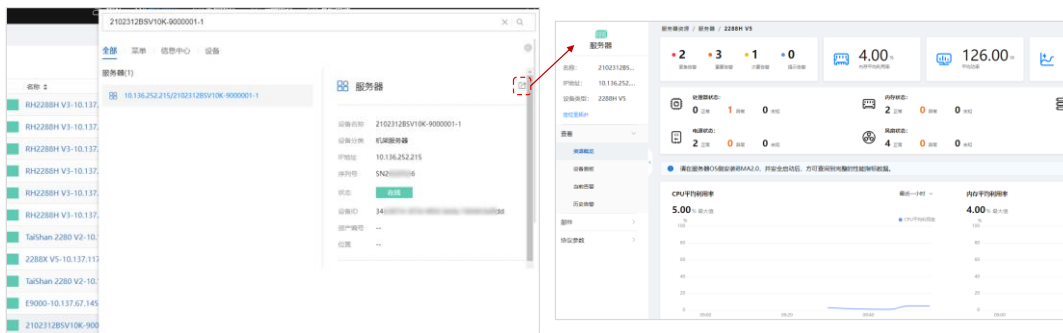
搜索		策略名称/设备名称/设备IP地址		高级搜索	实时数据	历史数据	性能策略
策略名称	设备名称	内存平均利用率	设备功率	平均功率	峰值功率	操作	
>	RH2288H V3-10.137.251.67	服务器设备默认监控策略	70.00% 200.00 W	无数据	无数据	无数据	无数据
>	RH2288H V3-10.137.251.66	服务器设备默认监控策略	70.00% 200.00 W	无数据	无数据	无数据	无数据
>	RH2288H V3-10.137.251.68	服务器设备默认监控策略	70.00% 200.00 W	无数据	无数据	无数据	无数据
>	E9000-10.137.67.145	服务器设备默认监控策略	无数据 1664.00 W	无数据	无数据	无数据	无数据
>	2288X V5-10.137.117.174	服务器设备默认监控策略	3.00% 264.00 W	265.00 W	516.00 W	无数据	无数据
>	RH2288H V3-10.137.251.63	服务器设备默认监控策略	70.00% 200.00 W	无数据	无数据	无数据	无数据
>	RH2288H V3-10.137.251.64	服务器设备默认监控策略	70.00% 200.00 W	无数据	无数据	无数据	无数据
>	RH2288H V3-10.137.251.65	服务器设备默认监控策略	70.00% 200.00 W	无数据	无数据	无数据	无数据
>	TaiShan 2280 V2-10.137.1...	服务器设备默认监控策略	无数据 258.00 W	246.00 W	660.00 W	无数据	无数据
>	TaiShan 2280 V2-10.137.1...	服务器设备默认监控策略	无数据 288.00 W	258.00 W	714.00 W	无数据	无数据
总计: 17				10条/页	2		

全局搜索

便捷的搜索入口，支持根据服务器的ID、名称、IP、SN、资产编号、位置等信息进行搜索



根据异常资源关键属性快速定位，跳转到资源MOUI进行360°分析



设备上下电管理：设备列表

- 服务器设备接入网管后，支持上下电操作的设备会自动展示到设备列表中，主要包括服务器设备的状态、名称、电源状态等信息。

上下电管理

设备列表

列表中的设备每6小时从服务器管理、存储管理中同步，如果不一致，建议手工同步。

过滤

支持上下电设备与存储和服务器中数量不一致时，点击同步进行设备对账

对单个设备执行上下电操作

同步

状态	名称	型号	类型	IP地址	电源状态	操作
离线	10.137.67.145-blade_6	CH121 V3	服务器	1.1.1.1	下电	⚙️
在线	Huawei.Storage	5500 V3	信息		上电	⚙️
离线	10.136.155.211-blade_2	CH121 V5	信息		上电	⚙️
离线	10.136.155.211-blade_4	CH121 V5	信息		下电	⚙️
离线	10.136.155.211-blade_6	CH121 V5	信息		上电	⚙️
离线	10.136.155.211-blade_1	CH121 V3	服务器	10.136.155.222	上电	⚙️
离线	10.136.155.211-blade_3	CH121 V3	服务器	10.136.155.223	上电	⚙️
离线	10.136.155.211-blade_5	CH121 V3	服务器	10.136.155.224	上电	⚙️
在线	RH588SH V3-10.136.1...	RH588SH V3	服务器	10.136.155.67	上电	⚙️
在线	Z288H V5-10.136.250...	Z288H V5	服务器	10.136.250.197	上电	⚙️

设备上电管理：任务管理

任务管理

- 创建任务对设备进行批量上下电操作
- 修改上下电任务
- 删除、恢复、挂起已创建任务



- 约束：服务器上下电状态实时更新依赖设备告警上报，需保证设备告警上报功能可用。否则需要在“服务器设备管理”界面单击“同步”获取最新的上下电状态。

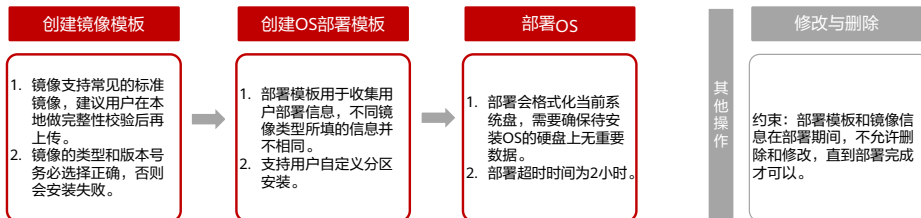
OS部署

场景：适合于开局、运维过程中的带内OS安装，支持主流操作系统。

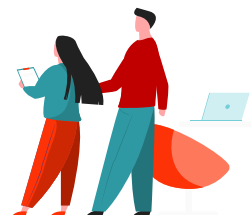
关键特性介绍：

- 1 镜像文件管理：管理系统中的OS镜像，提供增删改查功能。同时支持查看镜像空间的使用情况。
- 2 OS部署模板管理：管理系统中的部署模板，提供增删改查功能。模板支持**root用户名密码、OS语言、键盘类型、时区、安装位置分区**等通用参数，可选择是否部署iBMA。
- 3 执行部署：OS部署模板添加待部署设备后，可下发部署任务。部署任务支持不同设备差异化配置：**主机名**。

规格：每个任务支持**200台**设备，支持并发部署**40台**设备；当前仅支持含有Smart Provisioning的服务器。



OS部署示例



软件/固件升级

场景：适合于运维过程中的设备版本维护，支持固件、软件、驱动升级。

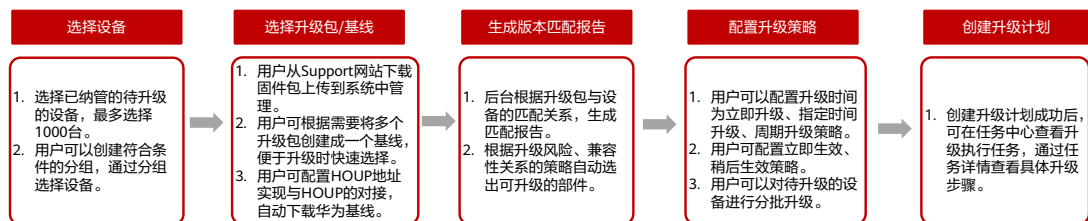
关键特性介绍：

- 1 升级包管理：支持用户手动上传升级包、对接HOUP下载华为基线、基于升级包创建自定义基线，通过基线管理快速获取需要的多个升级包。
- 2 设备版本管理：展示纳管设备所有部件和软件的版本信息，并提供软件/固件生效功能。
- 3 升级计划管理：选择设备和需要升级的软件/固件包，配置升级策略，包括配置升级时间、生效策略、分批策略，创建升级计划执行升级。

周期计划：选择分组和自定义基线创建周期计划，解决满足分组条件的新设备纳管后，可以自动升级到指定版本。

分批执行：对设备进行分批升级，可以灵活控制，确保第一批升级成功后，再升级下一批次。

规格：当前最多支持同时部署100台设备，一次升级计划最多选择1000台设备。



固件升级示例（1）



固件升级示例（2）




添加基线

基线名称: 请输入关键字

基线名称	类型	升级物数量	创建时间
✓ Huawei_latestBaseline	在线基线	280	
✓ 2	在线基线	280	
✓ 2	在线基线	77	
✓ 2	在线基线	255	
✓ 2	在线基线	173	
✓ 2	在线基线	51	
✓ 2	在线基线	185	

已选择基线: Huawei_latestBaseline

添加基线

设置执行策略

计划执行策略

☐ 立即执行

☒ 定时执行

计划执行策略: ☐ 每周 ☐ 每月 ☒ 按周时间

时间: 请选择日期时间

脚本生效策略

☒ 脚本生效

脚本生效策略: 脚本生效策略生效, 或在设备下次启动时生效。

☐ 立即生效

脚本生效策略: 脚本生效策略生效, 或在设备下次启动时生效。

脚本生效策略: 脚本生效策略生效, 或在设备下次启动时生效。

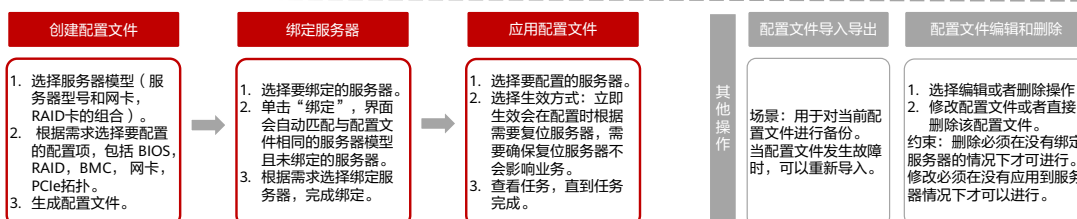
Profile配置部署

场景：将一台设备的相关配置形成一个Profile配置管理文件，进行批量配置，同时配置文件可导入、导出，可以快速复制到其它设备。

关键特性介绍：

- 1 服务器配置模型：查看已有服务器模型，可通过修改设备模型详细信息，使其与待规划的管理设备匹配。提高规划和管理效率。
- 2 服务器配置文件：主要为设备配置BIOS、RAID、BMC等参数。当配置文件应用于相应设备后，设备会按照配置参数进行配置。

规格：当前最多支持单个配置文件同时绑定2000台设备，一次应用最多选择1000台设备进行下发。



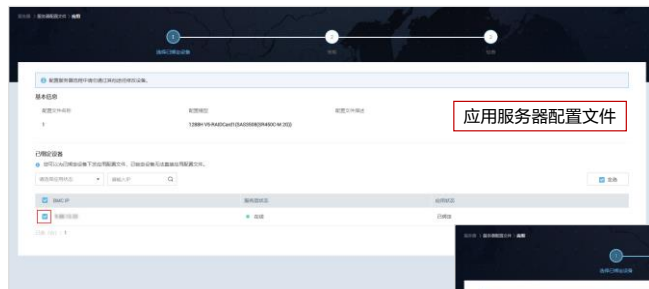
Profile配置部署（1）

创建服务器配置文件



绑定服务器

Profile配置部署（2）



目录

1. 计算运维软件体系
2. 单机管理
3. 集中管理
- 4. 近端运维**
5. 远端运维

SmartKit：华为IT产品近端运维作业工具箱



平台化、场景化、向导式是SmartKit工具箱的主要特点

- ✓ **一站式：**支持计算硬件产品（FusionServer、TaiShan、Atlas）整个生命周期的运维管理。
- ✓ **平台化、插件化、可扩展：**功能以插件的形式集成进SmartKit平台，插件支持独立更新，轻松扩展和升级。
- ✓ **向导式：**提供向导式界面，操作简单。
- ✓ **批量：**多设备并发处理，提升运维效率。
- ✓ **轻量化：**可安装于工程师笔记本，资源消耗低。
- ✓ **可定制：**支持白牌定制，支持接口定制后的产品。
- ✓ **在线更新：**支持平台和插件的在线更新。

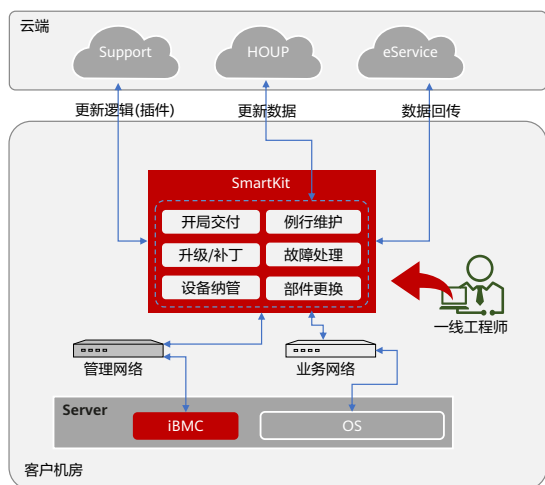
- SmartKit是华为IT产品的近端运维工具箱，当前已经是我们IT服务工程师笔记本必备的软件。
- SmartKit除了支持服务器，还支持华为的存储、机器视觉和云计算的产品。同时对于计算产品下的FusionServer、鲲鹏、Atlas等产品，提供了一站式运维能力。
- 对于开发者，SmartKit具备的平台化、插件化和可扩展能力，是的任何产品和团队都可以基于它开发新的功能。
- 对于用户，SmartKit具备的向导式、轻量化、在线更新等特性，可以大幅提高用户的交付和维护效率。
- 另外SmartKit的可定制能力，支持白牌定制，可以为鲲鹏等产品的生态伙伴快速构建运维能力。

SmartKit近端运维作业流程



- SmartKit近端运维大致分为以下几个步骤：
 - 进场前准备，升级工具到最新版本或指定版本，准备运维资料：包括升级要使用的软件包、模板、脚本和操作指导等。
 - 在获取客户授权后接入客户设备网络，如果需要，提前停止客户业务，然后添加设备。
 - 获取设备信息、配置任务参数。
 - 执行任务，业务恢复。

SmartKit应用场景



打造业界领先的近端运维工具，
提升服务/维护工程师现场交付和维护效率。

轻量

易安装

高效

并行处理

支持安装于普通PC，安装部署简单，
运行占用资源少。

上至50的并发处理能力，高效交付和
维护，减少服务工程师进站时间。

简单

向导式

兼容

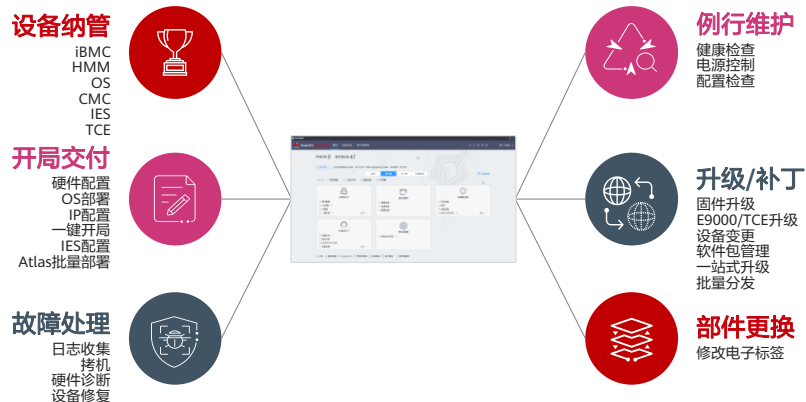
一站式

向导式界面操作，所见即所得，学
习成本低，技能要求低。

兼容华为FusionServer、TaiShan、
Atlas全系列产品，兼容主流新版本OS。

- 对于计算产品，SmartKit主要提供6大运维场景：开局交付、例行维护、升级、故障处理和部件更换。
- 各个场景又包括很多子功能，各个功能根据需要通过管理网络对接客户的iBMC或通过业务网络对接客户的设备的OS。
- 同时SmartKit支持对接云端，获取最新的运维资料，如最新的插件包、最新的软件包。或对接eService回传必要的运维数据。
- 为用户提供轻量、高效、简单和高可用的SmartKit，提高一线工程师现场交付和维护的效率。

SmartKit功能介绍



从IP初始化配置、OS部署到固件驱动升级和备件更换，
SmartKit提供了计算产品生命周期内运维所需要的所有功能。

- SmartKit的6大场景包含的功能：
 - 设备纳管代表了SmartKit支持的所有被管对象，基本所有的功能使用都需要先添加设备。
 - 开局交付包括配置部署。
 - 故障处理包括日志收集和诊断等。
 - 例行维护包括巡检和配置检查。
 - 升级包括固件升级、E9000升级等功能。
 - 部件更换包括电子标签修改等。

SmartKit常见工作界面



- SmartKit的场景工作界面：

- 使用SmartKit执行某项作业，首先进入的是场景向导界面。里面会列出完成该场景所需要执行的功能。
- 对于不同的功能会有不同的工作界面。

SmartKit功能介绍：IP配置

- 场景1：使用已知MAC地址或SN号为设备配置新的IPv4/IPv6/主机名。
- 场景2：使用已知的IPv4/IPv6地址为设备配置新的IPv4/IPv6/主机名。



- IP配置是针对机架类设备上架后，做IP初始化操作。
- 设备上架后的第一件是，就是修改设备的默认IP，修改方式有2种：
 - 通过已知的MAC地址或SN号来配置新的IP。
 - 通过已知的IPv4地址或IPv6地址为设备配置新的IP。

SmartKit功能介绍：IES配置

- 场景1：使用已知MAC地址或SN号为设备配置新的IPv4分区/证书等。
- 场景2：使用已知的IPv4/IPv6地址为设备配置新的IPv4/分区/证书等。



- IES配置是针对边缘小站Atlas500做的设备初始配置功能。
- 其初始化方案和IP配置类似，但其支持的配置项要比IP配置多很多。如IEF证书的导入和分区的配置。

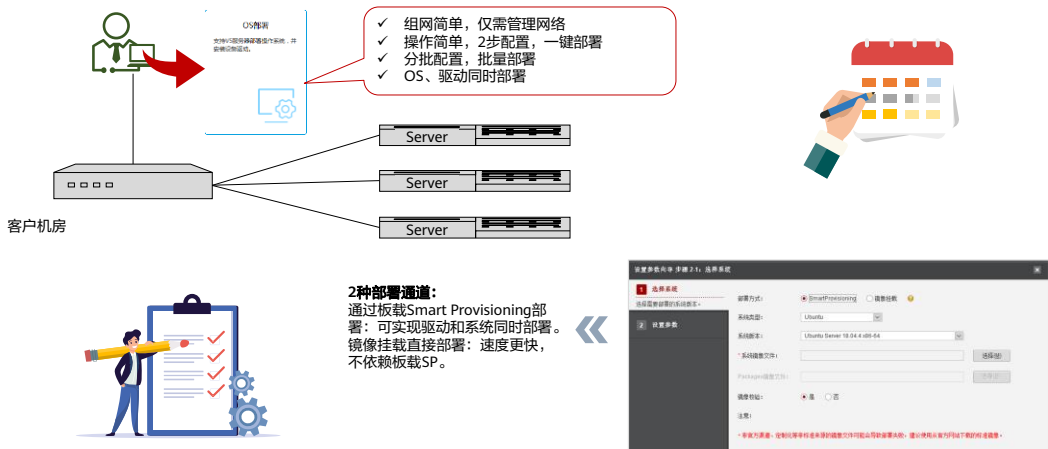
SmartKit功能介绍：硬件配置



- 初始IP修改好后，我们SmartKit就可以纳管设备了。接着就可以做开局配置部署。
- 硬件配置场景主要有4个步骤，配置、控制上下电生效和配置检查。
- 计算产品的硬件配置包括了iBMC配置、BIOS配置和RAID配置，对于E9000设备还有HMM的配置。

SmartKit功能介绍：OS部署

- OS部署：通过管理网络实现OS批量部署。



- 硬件配置完成后, 就具备了部署OS的条件,
- 客户自己做OS部署一般都使用PXE的方案, 对于客户自己, OS类型有限的场景下是没问题的, 搭建一套PXE部署服务器就可以一直用。
- 但是对于一线工程师来说, 不同的客户部署的OS类型、版本都不一样, 每进1个站点搭建1套PXE服务器是不现实的。
- 另外PXE部署方式需要通过业务网络来实现, 这也增加了PXE 使用的门槛。
- SmartKit提供的OS部署功能支持通过管理网络批量部署OS。只需要SmartKit通过交换机批量接入管理网络即可。
- 后台实现上, SmartKit提供了2种部署方式。
- 通过Smart Provisioning部署可以同时安装SP里带的驱动。
- 通过镜像挂载部署, 可以不依赖SP, 但是需要在安装好系统后额外安装驱动。

SmartKit功能介绍：一键开局

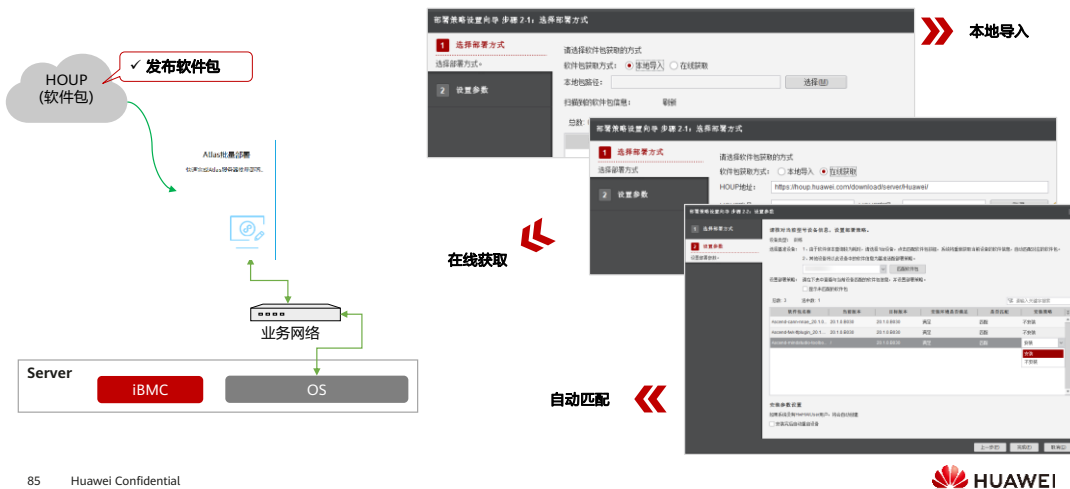
- 一键开局：整合配置部署流程，简化部署升级操作，缩短开局交付操作工时。



- 开局通常都需要对iBMC、BIOS和RAID进行配置，前面讲的配置部署功能都是独立的，每个功能执行都需要走一遍向导，经历1设备信息获取、参数配置和等待执行。
- 一键开局整合了配置部署功能，也支持用户自定义配置流程，支持模板导入。

SmartKit功能介绍：Atlas批量部署

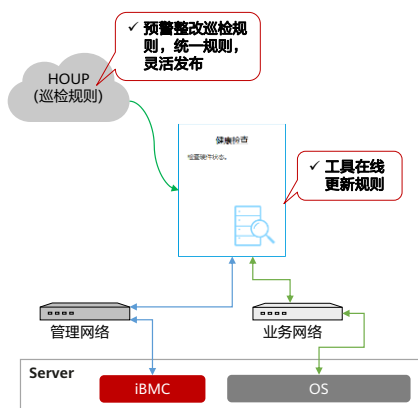
- 对接OS，为Atlas系列产品部署CANN软件、向网管注册。



- 交付场景的最后1个功能是Atlas批量部署。此功能主要是为Atlas系列产品部署CANN软件。
- 该功能对接的是设备的业务面。通过SSH协议向设备传输和安装CANN软件包，当然该功能后续也支持从HOUP下载软件包，直接向设备部署，不需要用户去网站下载。

SmartKit功能介绍：健康检查

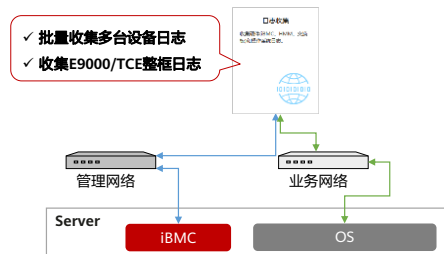
- 对接OS和iBMC，对设备进行批量巡检，快速排查设备风险。



- 例行维护里最常用的功能就是健康检查了，在uMate时代，我们叫巡检，收编到 SmartKit后，改名叫健康检查。
- 健康检查主要提供很多事项等检查项。检查完后输出检查报告。
- 健康检查支持带内和带外，在带内网络和带外网络没有隔离的情况下可以一起检查。
- 另外实现了部分检查规则和工具的解耦，主要是预警整改规则上HOUP。

SmartKit功能介绍：日志收集

- 多场景日志收集方案，收集设备带内带外日志。



OS日志收集脚本包

Huawei SmartKit\tools\ServerInspect\tools\InfoCollect\tools
FusionServer Tools-SmartKit-ServerInspect-V170.zip\tools\InfoCollect\tools\

名称	修改日期	类型	大小
InfoCollect_Linux.tar.gz	2023/11/11 1:04	gz Archive	25,043 KB
infoCollect_Windows.zip	2023/11/11 1:04	zip Archive	12,763 KB

获取设备信息

☒ 一键日志收集 ☐ E9000自定义日志收集 ☐ TCE自定义日志收集

☒	IP地址	用户名	设备类型	序列号	产品名称	厂商
☒	71.46.142.249	root	HMM	21030546_	MM910	

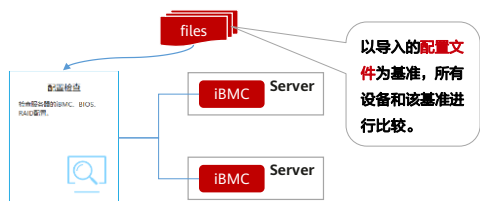
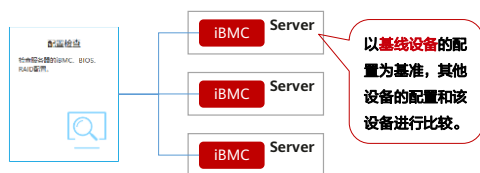
多种收集场景
一键日志收集
E9000自定义日志收集
TCE自定义日志收集

E9000整框收集
管理板
计算节点BMC
计算节点OS
交换板

- 例行维护的另1个重要功能是日志收集。
- 日志收集分几种常见：
- 最常用的场景一键日志收集：批量收多台。
- E9000和TCE的自定义日志收集，可以针对E9000/TCE整框进行收集，可以按需收集交换板各平面日志。收集的深度更深。
- 日志收集也支持带内和带外收，依据客户提供是IP、账号和密码自动识别设备类型进行收集。
- 在带内日志收集场景下，如果客户不同意我们接入其业务网络，我们提供的OS收集叫本包。

SmartKit功能介绍：配置检查

- 从BMC获取设备配置信息，和基准进行比较，识别配置差异。



您可以导入标准配置文件或选择待查设备其中之一作为iBMC、BIOS、RAID的校验标准。

☒ 检查iBMC配置

☒ 选择标准设备: 71.44.1.136(2288H V5) 确定

☐ 导入标准配置文件: 选择 清除

☒ 检查BIOS配置

☒ 选择标准设备: 71.44.1.136(2288H V5) 确定

☐ 导入标准配置文件: 选择 清除

☐ 导入模板文件: 选择 清除

☒ 检查RAID配置

☒ 选择标准设备: 71.44.1.136(2288H V5) 确定

☐ 导入标准配置文件: 选择 清除

设备和配置文件2种对比方式

多维度查看对比报告

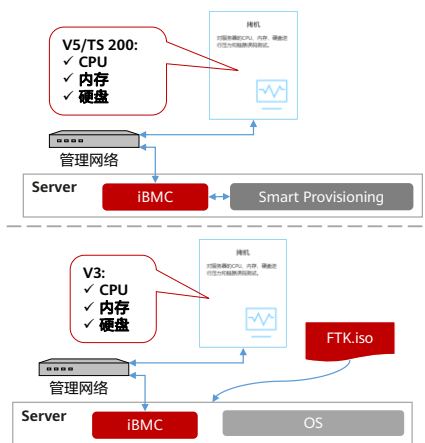
标准配置 导出配置对比汇总表 导出iBMC配置项差异详情 导出BIOS配置项差异详情 导出RAID配置项差异详情

配置信息对比表	配置项差异表	一致	不一致	总计
71.44.1.137 2288H V5 iBMC	549	75	620	
71.44.1.137 2288H V5 BIOS	757	4	761	
71.44.1.137 2288H V5 RAID	198	0	198	

- 例行维护的另外一个功能是配置检查，是通过和基准设备和基线文件对比来排查设备间的配置差异。提供了2种对比方式：
 - 以配置正确的设备作为基准设备进行比较。
 - 以基线文件为基准，所有设备和基线文件进行比较。基线文件可以通过该功能导出，留存下次进展时使用。

SmartKit功能介绍：拷机

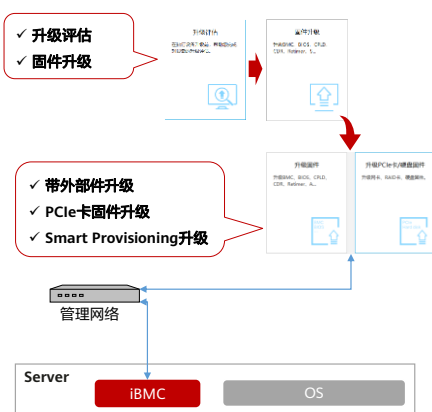
- 通过压力测试，排查三大件（CPU/内存/硬盘）故障。



- 下面进入故障诊断场景，常用功能是拷机。主要应用场景是设备上架后批量拷机，排查设备运输过程中产生的硬件故障。拷机的主要原理是对三大件进行压力和长时间测试，识别设备问题。
- V5、TS200的设备通过iBMC下发命令给Smart Provisioning，有SP来执行压力测试，V3的设备通过iBMC挂载FTK来进行压测。压测完后会输出压测报告。

SmartKit功能介绍：固件升级

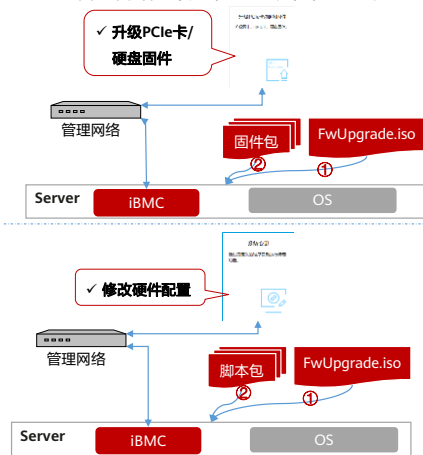
- 通过iBMC升级带外固件和带内（PCIe卡/硬盘）固件。



- 服务器产品里需要升级的部件有iBMC、BIOS、SP、PCIe卡固件、硬盘固件等。其中BMC、BIOS和SP属于带外部件，通过iBMC可以直接升。而PCIe卡和硬盘固件我们称为带内固件，我们升级一般要进入业务系统或者外挂内存小系统。带外固件升级提供了单部件升级和多部件升级。

SmartKit功能介绍：PCIe固件升级/设备变更

- 通过管理网络挂载内存小系统，在小系统内执行变更脚本。



通过挂载FwUpgrade对设备进行变更操作

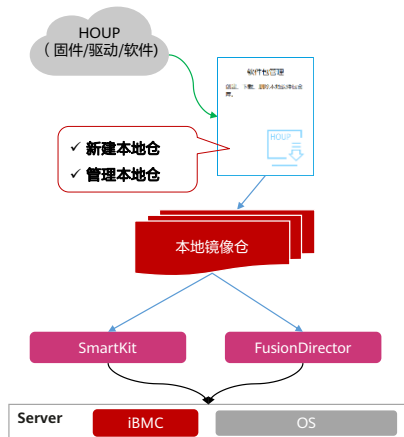
通过挂载FwUpgrade升级PCIe卡/硬盘固件



- PCIe卡和硬盘固件升级和设备变更的原理是一样的。V5设备已经提供了接口，可通过iBMC传包和下发命令给SP，设备变更：挂载内存操作系统，执行自定义脚本。

SmartKit功能介绍：软件包管理

- 软件包管理：提供本地HOUP镜像仓的创建和管理能力。



软件包管理

新建本地仓

从HOUP平台创建仓库

管理本地仓

删除、转移本地仓

软件包管理

当前需要安装10个版本

序号	版本号	版本	大小	更新时间	已下载/总数	操作
1	20200624	2020.06.24	16.31GB	2020-06-24 12:14:48	- / -	下载
2	20200602	2020.06.02	1.79GB	2020-06-02 16:19:23	6 / 76	下载
3	20200618	2020.06.18	16.31GB	2020-06-18 14:05:43	6 / 198	下载

本地镜像仓列表

仓库名	产品名	Module	版本	数据源地址	大小	状态	备注
>	Z288H V5A128...	Z288H V5	CPUD	2.12	NA	151.41KB	未下载
>	Z288H V5A228...	Z288H V5	CPUD	1.11	NA	20.09KB	未下载
>	Z288H V5A228...	Z288H V5	CPUD	1.12	NA	20.09KB	未下载
>	Z288H V5A228...	Z288H V5	CPUD	1.11	NA	20.09KB	未下载

新建本地仓

管理本地仓

下载固件

下载驱动

下载软件

- 软件包管理主要是解决网管和HOUP无法直通的问题，他提供了新建和管理本地仓的功能。

SmartKit功能介绍：一站式升级

- 一站式升级：对接HOUP下载固件，一键升级多个部件。

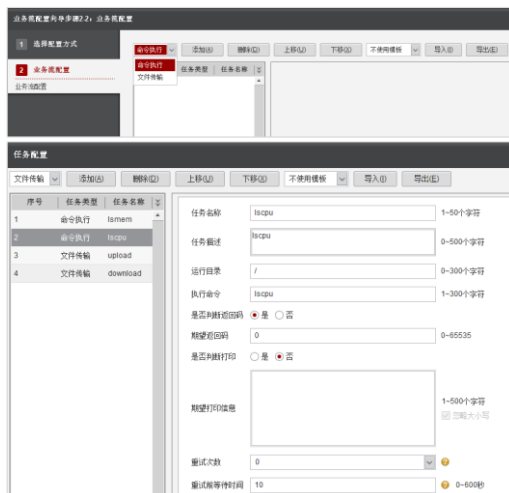
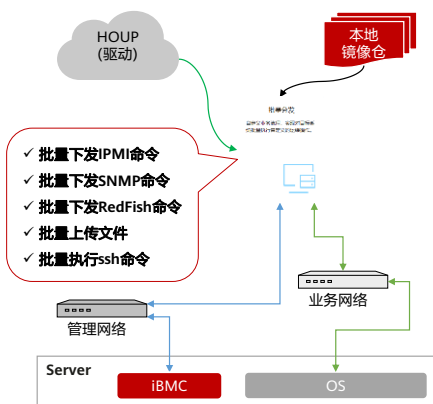


- ✓ 一站式升级
- ✓ 本地导入固件包
- ✓ 在线下载固件包
- ✓ 使用本地镜像仓

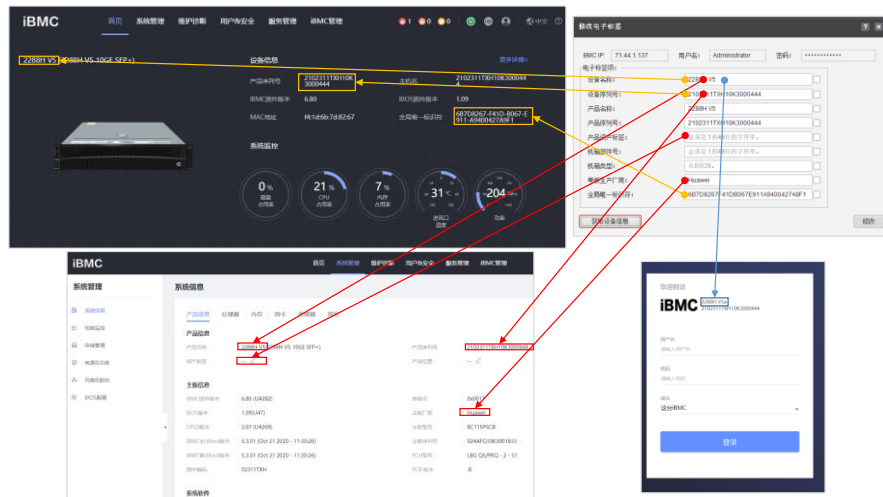
- 一站式升级，类似于一键开局，整合了包下载和升级的流程。
- 可以一站升级带内和带外固件。
- 不同设备可以设置不同的升级策略。

SmartKit功能介绍：批量分发

- 批量分发：向BMC、OS批量分发命令、脚本和软件包。



电子标签项和修改目标的对应关系

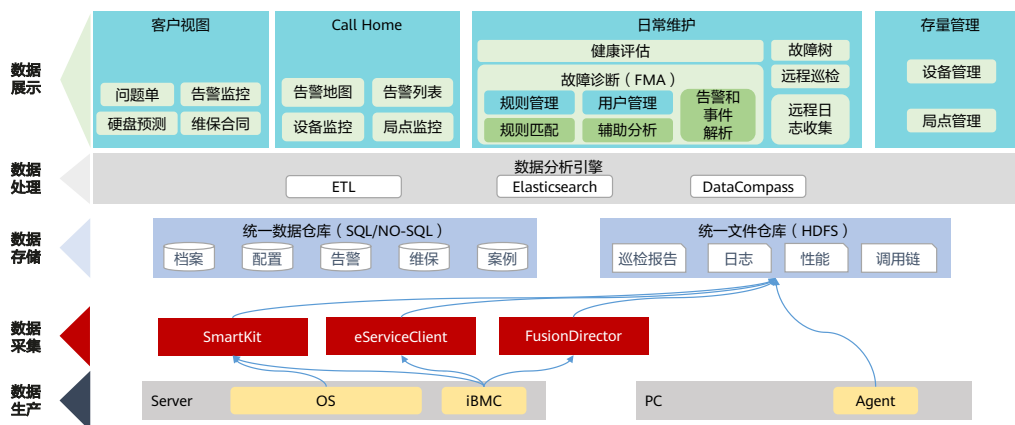


目录

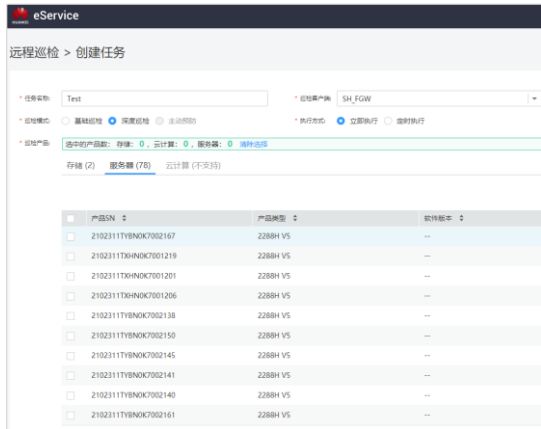
1. 计算运维软件体系
2. 单机管理
3. 集中管理
4. 近端运维
5. **远端运维**

eService: 远程运维平台

eService作为计算领域的远程运维平台，集中管理计算产品现网的各种运维基础数据，聚焦“维护和存量经营”两个价值场景：提升故障处理效率，支撑产品自动化运维能力的构建；实现服务变现，提升客户满意度。



eService功能介绍：远程巡检/日志收集



远程巡检 > 创建任务

任务名称: Text

任务来源: SH_FGW

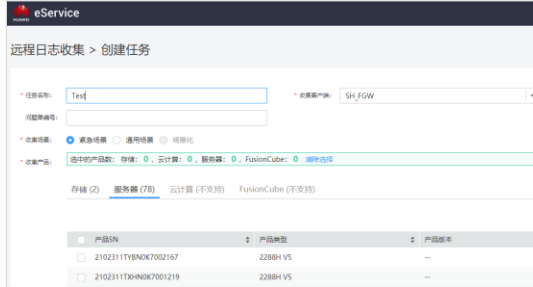
巡检模式: ☐ 基础巡检 ☒ 深度巡检 ☐ 主动巡检

执行策略: ☒ 立即执行 ☐ 定时执行

选中产品数: 存储: 0, 云计算: 0, 服务器: 0, FusionCube: 0, 网络存储: 0

存储 (2) 服务器 (78) 云计算 (不支持) FusionCube (不支持)

产品SN	产品类型	软件版本
☐ 2102311TYBNOK7002167	2288H V5	...
☐ 2102311TXH-NOK7001219	2288H V5	...
☐ 2102311TYBNOK7001201	2288H V5	...
☐ 2102311TXH-NOK7001206	2288H V5	...
☐ 2102311TYBNOK7002138	2288H V5	...
☐ 2102311TYBNOK7002150	2288H V5	...
☐ 2102311TYBNOK7002145	2288H V5	...
☐ 2102311TYBNOK7002141	2288H V5	...
☐ 2102311TYBNOK7002140	2288H V5	...
☐ 2102311TYBNOK7002161	2288H V5	...



远程日志收集 > 创建任务

任务名称: Text

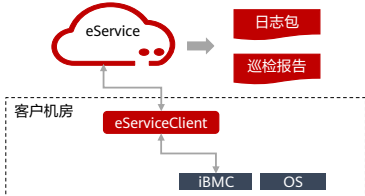
任务来源: SH_FGW

收集策略: ☒ 基础收集 ☐ 通用收集 ☐ 高级收集

选中产品数: 存储: 0, 云计算: 0, 服务器: 0, FusionCube: 0, 网络存储: 0

存储 (2) 服务器 (78) 云计算 (不支持) FusionCube (不支持)

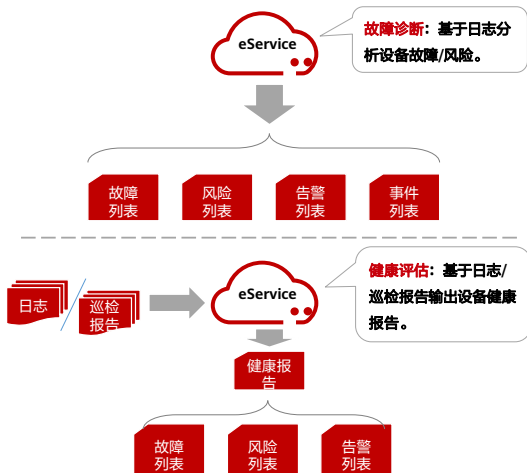
产品SN	产品类型	产品版本
☐ 2102311TYBNOK7002167	2288H V5	...
☐ 2102311TXH-NOK7001219	2288H V5	...



```

graph TD
    eService((eService)) --> LogPkg[日志包]
    eService --> Report[巡检报告]
    subgraph CustomerDC [客户机房]
        eServiceClient[eServiceClient]
        iBMC[iBMC]
        OS[OS]
        eServiceClient --> iBMC
        eServiceClient --> OS
    end
    eServiceClient --> eService
    
```

eService功能介绍：故障诊断/健康评估



新建分析

选择设备: 智能服务器

选择分析组: ☒ 问题匹配 ☒ 辅助分析 ☒ BMC管理端事件

优先导入文件:

选择规则: 所有规则

设置型号:

任务类型:

选择文件:

创建健康评估任务

任务基本信息

任务名称:

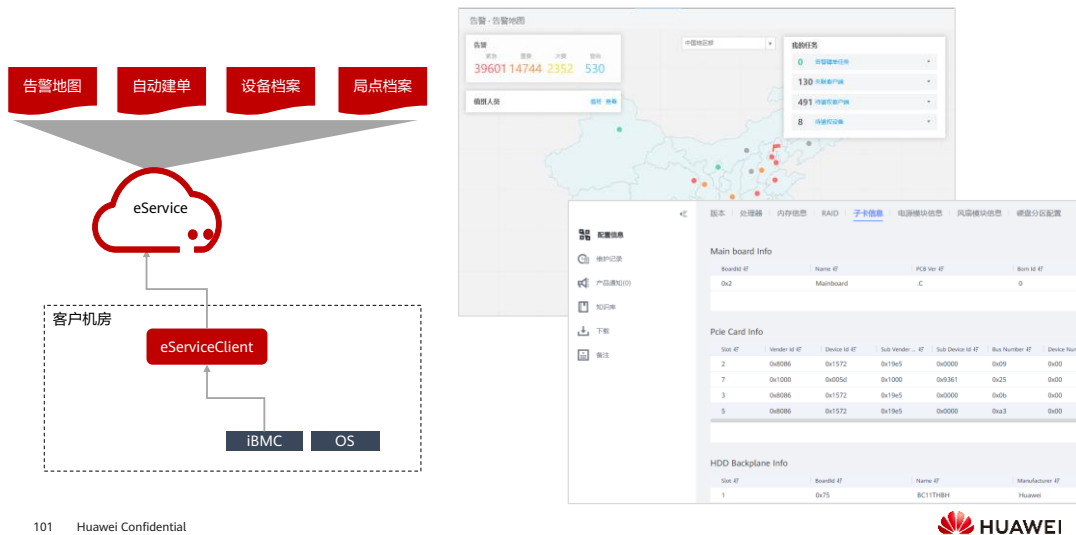
设备名称:

工单号:

报告类型: ☒ 健康评估报告 ☐ 离境巡检报告

导入数据包 支持导入压缩包(.zip),
1. 建议用SmartKit V2R5C00R07及以上版本工具生成;
2. 压缩包内目录下包含*.xls的巡检文件。

eService功能介绍：Call Home



- Call Home是把华为设备产生的硬件告警回传到华为支持中心，并创建问题单，进行快速处理，是维保服务新增的一种自动报障手段。

思考题

1. （多选题）某局点采购一批TaiShan 200系列2280均衡型服务器，这批服务器并没有预集成软件。工程师在进行项目交付时，需要对这批服务器进行RAID配置，并安装操作系统，以下哪些计算管理软件可以进行该型号服务器型号的RAID配置？
- A. Smart Provisioning
 - B. SmartKit
 - C. iBMC
 - D. eService

- 参考答案：ABC

本章总结

- 本章主要讲述了：
 - 计算产品运维软件体系
 - 单机带内管理iBMA的功能
 - 单机带外管理iBMC、Smart Provisioning的功能
 - 集中管理FusionDirector的功能
 - 集中管理eSight的功能
 - 近端运维SmartKit的功能
 - 远端运维eService的功能

学习推荐

- 智能管理软件 文档中心：
 - <https://support.huawei.com/enterprise/zh/category/management-software-pid-1548148290314?submodel=doc>

缩略语

- GPU: Graphics Processing Unit, 图形处理器, 一种专门在个人电脑、工作站、游戏机等设备上图像运算工作的微处理器, 可用于桌面的图片渲染。
- NPU: Neural Processing Unit, 网络处理单元, 一种AI芯片, 通过承担小的重复性进程来减少GPU和CPU上的一些负载, 使计算机在满足AI驱动的请求时可以更高效地工作。
- FPGA: Field Programmable Gate Array, 现场可编程门阵列, 专用集成电路领域中的一种半定制电路。
- CANN: Compute Architecture for Neural Networks, 华为针对AI场景推出的异构计算架构, 通过提供多层次的编程接口, 支持快速构建基于Ascend平台的AI应用和业务。
- HOUP: Huawei Online Upgrade Repository, 华为在线升级仓库。
- TCE: Telco Cloud Engine, 云核心网投资开发的下一代硬件平台, 采用4U/8U机框, 支持框内计算刀片共享硬件管理、风扇、电源。
- IES: Intelligent Edge Site, 智能边缘小站, 将云基础设施和云服务部署到企业现场。

缩略语

- HMM: Hyper Management Module, 超级管理模块, 负责电源和风扇的集中管理, 是服务器节点iBMC的管理延伸, 主要负责电源和风扇健康状态监控、资产信息和实时数据获取、散热调速、网络汇聚。
- FTK: FusionServer Tools Toolkit, 服务器工具集, 主要用于服务器运维。
- IB: InfiniBand, 一种交换结构I/O技术。
- RTOS: Real-time Operating System, 当外界事件或数据产生时, 能够接受并以足够快的速度予以处理, 其处理的结果又能在规定的时间之内来控制生产过程或对处理系统做出快速响应, 调度一切可利用的资源完成实时任务, 并控制所有实时任务协调一致运行的操作系统。
- ETL: Extract-Transform-Load, 抽取转换与加载, 常用在数据仓储。
- OPEX: Operational Expenditure, 运营支出。
- PON: Passive Optical Network, 无源光网络, 宽带接入光纤技术。
- OTA: Over the Air Activation, 空中激活, 通过移动通信的空中接口实现对移动终端远程管理的技术。

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。
Bring digital to every person, home, and
organization for a fully connected,
intelligent world.

Copyright©2024 Huawei Technologies Co., Ltd.
All Rights Reserved.

The information in this document may contain predictive statements including, without limitation, statements regarding the future financial and operating results, future product portfolio, new technology, etc. There are a number of factors that could cause actual results and developments to differ materially from those expressed or implied in the predictive statements. Therefore, such information is provided for reference purpose only and constitutes neither an offer nor an acceptance. Huawei may change the information at any time without notice.



专业技术人员职业素养与组织赋能



前言

- 在技术高歌猛进、业务复杂多元的今天，我们坚信，卓越的专业追求绝非孤立存在，它既需要**内化于心的职业道德**——对专业诚信的坚守、对合规红线的敬畏、对技术伦理的审慎；也需要**外化于行的组织赋能**——打破技术孤岛，通过知识沉淀、经验传承，将个人能量转化为组织发展的动能。

目标

- 学完本课程后，您将能够：
 - 清晰界定技术应用的边界，识别风险；
 - 建立持续迭代的技术成长思维，以适应指数级发展趋势；
 - 掌握标准化带教方法，将个人技术传承延续。

目录

1. 职业道德
2. 终身学习
3. 培训指导

从业者的三重责任

三重责任

核心目标：明确责任边界与核心义务

1. **专业责任：**追求技术卓越，保持专业精进
2. **社会责任：**维护公共利益，坚守伦理底线
3. **带教责任：**传承知识与规范，赋能团队成长



三者间的关系

职业道德 → 技术技能（约束关系）

职业道德为技术技能的应用划定边界。如生成式AI伦理约束AI模型开发方向；网络安全法规约束网络技术操作

培训指导 → 技术技能（赋能关系）

培训指导为技术技能的习得提供方法论。如利用知识图谱工具构建能力图谱指导学习路径；“先培训再上岗”确保技术技能达标

技术技能 → 职业道德（实践关系）

技术技能是践行职业道德的载体。如只有掌握了数据安全技术，才能真正保护用户隐私

技术技能 → 培训指导（反馈关系）

技术技能的掌握程度反馈培训指导的有效性，驱动培训方案迭代

职业道德 ↔ 培训指导（融合关系）

培训指导本身需遵循职业道德；同时培训指导也是传播职业道德的主渠道

职业道德核心原则解析

一般伦理原则

核心目标：确立社会底线

1. **人类福祉**：将公众安全、健康及福祉置于技术发展的首位
2. **诚实守信**：保持专业诚信，不隐瞒已知风险，不误导用户
3. **公平公正**：避免算法歧视，确保技术应用的包容性与公平性
4. **避免伤害**：预防技术滥用，主动规避对社会的潜在负面影响
5. **尊重权益**：尊重法律、知识产权以及用户隐私，严守数据边界

职业伦理原则

核心目标：确立行业标准

1. **追求卓越**：持续精进技术能力，保持对高质量交付的追求
2. **保持专业**：遵循行业规范，维护职业声誉，保持客观中立
3. **同行评议**：确保研究成果的质量和可信度，维护学术和职业的公平和正义
4. **尽职尽责**：对岗位负责，确保系统稳定与安全
5. **维护专业形象**：维护计算机行业的形象与声誉，促进计算机技术的普及及推广

一般伦理原则与职业伦理原则的双重规范

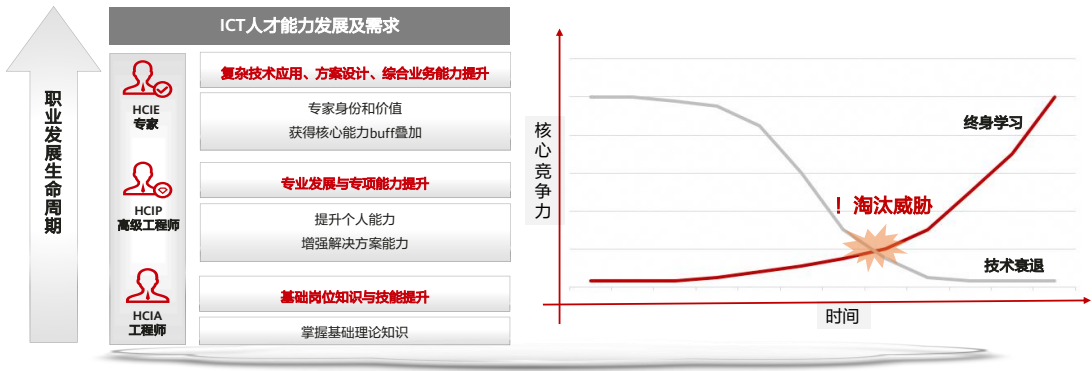
目录

1. 职业道德
- 2. 终身学习**
3. 培训指导

数智升级，对人才培养与技能评价提出更高要求



华为职业认证：终身学习



技术发展现状

- 迭代加速、跨界融合、技术颠覆成为常态，停止学习即被淘汰
- 从“碎片化输入”到“知识晶体”，真正的学习，是将外界的碎片重构成自身的“护城河”

目录

1. 职业道德
2. 终身学习
- 3. 培训指导**

角色升级：从技术能手到培训导师

“技术工作是把事做好，培训带教是把人带会”

	技术能手	培训导师
核心角色	解决特定技术问题	知识传递者、能力引导者、问题解决者
交付成果	高质量的代码与系统	独立规范上岗的专业人才
核心能力	专业硬技术、排障能力	会梳理知识、会标准示范、会纠错复盘



带教原则

实用为先

贴合岗位真实场景，不讲空话

实操为主

重动手，重细节，真操实练，夯实技术

问题导向

解决实操卡点与上岗痛点

全程闭环

课前有备

定目标，梳理重难点

课中有方

示范、纠错、控场、盯安全

课后有落

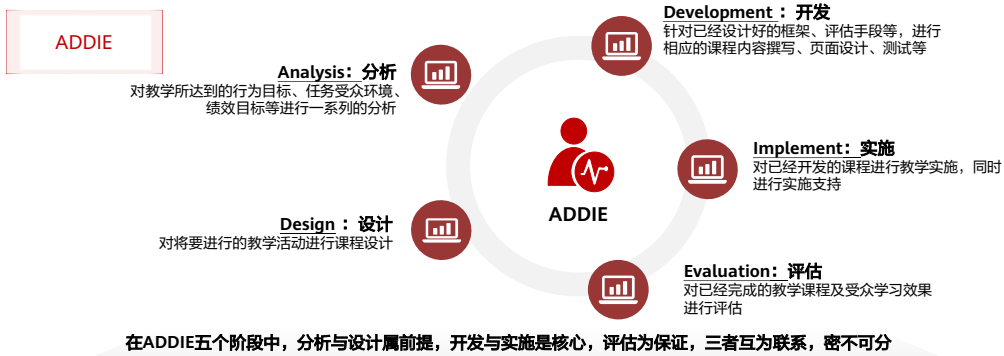
答疑、考核、建档、跟进成长

培训与指导

ADDIE：将A、D、D、I、E五个步骤综合起来考虑，形成系统，保障各个环节能进行及时有效的评估

SAM模型：提倡“快速迭代、小步快跑”，分为三阶段：准备、迭代设计、迭代开发

6D法则：以终为始，把培训的生命周期向后延伸，将“培训”转化为“商业结果”的全面流程



Analysis: 分析阶段

分析阶段是首要的环节，是确定培训目标、设计培训计划的前提，也是决定培训是否开展的关键因素

在培训的各个阶段都要有分析，通过问卷、约谈、电话等各种形式，了解对象、组织、课程的各方面需求



需求分析 (WHY/WHAT) :

组织需求 (公司层面) :

公司的战略目标是什么?
目前阻碍目标实现的“痛点”是什么?

岗位需求 (标准层面) :

目标岗位需要具备哪些核心胜任力?
标准的操作流程 (SOP) 或优秀员工的行为典范是怎样的?



需求对象分析 (WHO) :

基本属性 (画像) :

学员的年龄分布、学历、司龄、目前的岗位层级。

认知与经验起点:

他们对这个主题已经了解多少?
他们日常遇到的实际工作场景和痛点是什么?

学习动机与风格:

他们是主动想学，还是被公司强迫来的?



培训环境分析 (Where & How)

面授课堂: 教材需要配套 PPT、讲师手册、学员手册、通关演练通关单。

线上微课/E-learning: 教材需要拆解为结构化的“脚本”或“脚本大纲”，方便录制音频或制作动画。

混合式培训 (OMO) : 线上学理论 (教材偏精简、清单化)，线下做演练 (教材偏案例、工具化)。



培训内容分析 (Which)

知识类: 讲概念、原理 (要少而精)。

技能类: 讲步骤、方法、流程 (要多、要细、要有标准)。

态度类: 讲观念转变、企业文化 (要通过故事和案例来感染)。



培训资源分析 (With What)

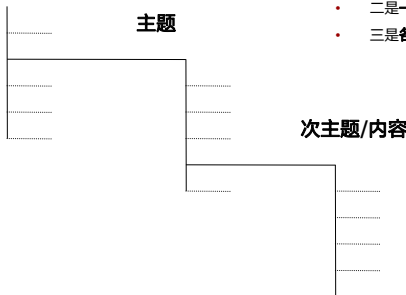
师资资源: 有没有内部的业务专家 (SME) 作为内容源头?

素材资源 (现成原料) : 公司现有的 SOP 手册、过往的失败案例库、竞品资料、行业标准等。

Design：设计阶段

设计阶段要制订一份完成培训项目所必需的说明提纲。基本内容包括必须制定的目标、目的和评估任务
基于培训项目的相关决策和其它实施问题的体系结构，确定各项任务的顺序

单元



课程单元结构的关系

- 一是**相对独立型**，各学习单元在位置上可以互换
- 二是一个单元是构成另一个单元的基础，这种类型在序列上极为严密
- 三是各单元学习内容的联系成**综合型**，即有的是相互独立关系，也有的是严格序列关系

主题的组织方式 - 螺旋式

- 即根据培训对象的智力发展水平，让学员尽早有机会在不同程度上去接触和掌握某门学科的基本结构，以后随着学员在智力上的成熟，围绕基本结构不断加深内容深度，使学员对知识有更深刻和有意义的理解。

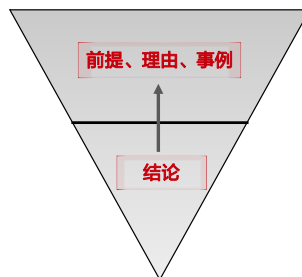
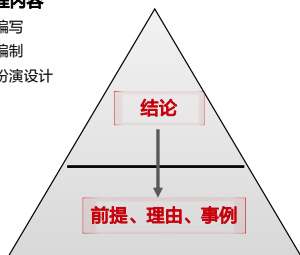
Development : 开发阶段

开发阶段的要点在于**开发创建教学材料**的方法和流程。根据之前总结的各干系方需求，开发培训教材（资料），设计好阶段性评估文件等。在这一阶段中，内容和材料从概念设计转变为可交付成果。

1: 收集素材

2: 组织课程内容

- 内容编写
- 案例编制
- 角色扮演设计



检查错误

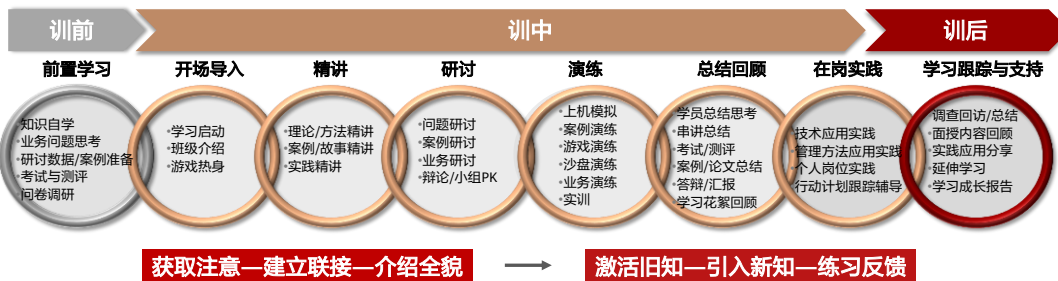
- 标题
- 段落
- 句子
- 表格
- 图形
- 缩略语/术语
- 数与数量
- 标点符号
- 单位符号

思考时多按金字塔顺序，但表述要用倒金字塔顺序

Implement: 实施阶段

- **教学程序**: 从把学员引入某个专题, 到掌握某个目标为止向学员呈现教学时教学活动的程序、步骤
 - **教学方式**: 呈现教学时运用的手段、方法, 如分组讨论、独自阅读、个案研究、讲授、计算机模拟、小组协作项目等
- 认知学习的一个关键概念是在信息从**工作记忆**迁移至**长期记忆**时, 对信息进行提取或演练, 从而强化长期记忆

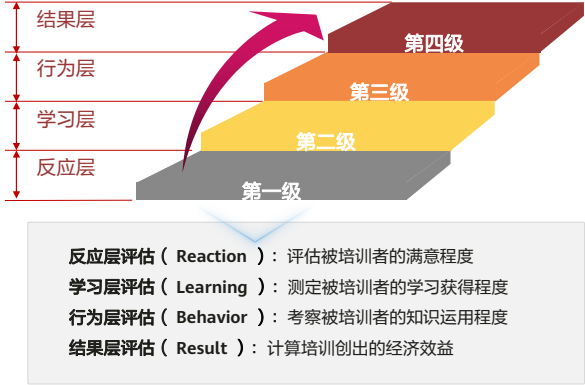
“三段八环六素”教学理论, 把学习过程分片, 根据不同技术场景, 适配不同教学活动, 提升学习效果。



Evaluation: 评估

主流评估模型：柯氏四级、菲利普斯五级、考夫曼五级、CIPP 模型等
行业地位：历经行业检验，**柯氏四级**时至今日仍是不可动摇的经典基石模型

层次	重点描述	衡量方法	评价人员
反应层	受训者满意程度：喜欢该项目吗？对培训人员和设施有什么意见吗？课程有用吗？	问卷	学员
学习层	受训者在培训前后，知识、技能、态度、行为方式方面的收获：有多大程度的提高？	笔试、实地操作和工作模拟	学员和培训人员
行为层	培训后，受训者行为的改进：行为有无不同？他们在工作是否使用了在培训中所学的知识？	问卷调查、面谈、直接观察、业绩检测、专题研究、实践计划	学员、学员上级、下属、同事、客户
结果层	受训者获得的经营业绩：组织是否因为培训变得更好？	问卷调查、业绩检测、业绩目标法	学员、学员上级



思考题

1. （单选题）项目经理小张突然接到紧急通知，需要当晚在公司内网加班处理一份涉及公司**核心商业秘密与关键技术参数**的涉密文件。由于公司保密机房的专用有线计算机键盘出了故障，为了不耽误进度，小张决定采取一些临时措施。请问以下哪**一项行为是明确禁止**的？
 - A. 临时将这份涉密文件拷贝到自己带有 Wi-Fi 芯片和蓝牙功能的个人便携式笔记本电脑上处理。
 - B. 立即联系IT安全部门，申请对当前的涉密系统环境进行一次全面的安全性评估。
 - C. 调整工位，移步到未链接互联网、且切断一切外网物理信号的涉密专用计算机上处理。
 - D. 确保待在公司规定的受控办公室内，在无无关人员围观的闭环环境下处理该涉密资料。
2. （判断题）培训完成后，进行了结业考试，考试通过就代表学员完全掌握了这项技能。

- A
- 错

本章总结

- 本章讲述了从业者的成长路径
 - 坚守底线：牢记行为红线，确保技术向善
 - 持续学习：践行终身学习，应对瞬息万变
 - 培训指导：应用教学实践，传承优秀同路人

学习推荐

- 《中国计算机学会职业伦理与行为守则》：
https://www.ccf.org.cn/About_CCF/CCF_Constitution/2023-07-26/794450.shtml

Thank you.

把数字世界带入每个人、每个家庭、
每个组织，构建万物互联的智能世界。
Bring digital to every person, home, and
organization for a fully connected,
intelligent world.

版权所有 © 华为技术有限公司 2026。

保留一切权利。

未经华为事先书面许可，任何人不得对本资料进行复制、修改、
整合、改编、翻译，也不得将本资料或其任何部分或基于本资
料的衍生作品提供、出售、转让、再许可给任何人用于华为
许可以外的任何其他用途。

Copyright © Huawei Technologies Co., Ltd. 2026.

All Rights Reserved.

Without prior written permission of Huawei, no one is
allowed to reproduce, modify, integrate, adapt, translate
this Document, or provide, sell, transfer, sublicense this
Document or any part of this Document or its derivative to
any third party for any other purpose otherwise permitted
by Huawei.

