



联想政教行业 HPC解决方案



手机专线 400 819 6776
座机专线 800 830 6776
dcg.lenovo.com.cn

Ultrabook、赛扬、Celeron Inside、Core Inside、英特尔、英特尔标志、英特尔凌动、Intel Atom Inside、英特尔酷睿、Intel Inside、Intel Inside 标志、英特尔博锐、安腾、Itanium Inside、奔腾、Pentium Inside、vPro Inside、至强、至强融核、Xeon Inside 和英特尔傲腾是英特尔公司或其子公司在美国和/或其他国家（地区）的商标。

咨询电话 400 819 6776

英特尔®至强®可扩展平台 加速探索与获取洞察

联想政府科研教育行业通用 HPC 解决方案

系统总体架构

联想高性能计算系统分为以下 6 个子系统：

计算系统：用于执行用户计算任务的子系统

根据计算系统中的服务器节点的硬件配置及提供的计算服务的不同，计算系统中的服务器节点可分为 CPU 计算节点、GPU 加速节点和大内存多核胖节点。



CPU计算节点



GPU加速节点



大内存多核胖节点

存储系统：用于存储用户数据或计算任务数据的子系统。根据存储系统的架构设计和构建。



存储节点与磁盘阵列



传统分布式存储系统



DSS-G高端并行存储系统

服务系统：用于管理和监控集群系统以及用户接入的子系统。根据提供的服务类型不同，服务系统中的功能节点分为登录节点、可视化节点、管理节点和监控节点四类。



登录节点

系统使用者通过该节点登录集群进行作业。



可视化节点
(包括远程可视化节点)

系统使用者远程可视化。



管理节点

系统管理员通过该节点管理和配置整个集群。



监控节点

系统管理员和运维人员通过该节点对整个集群系统的软硬件资源进行监控。

网络系统：用于子系统内连接和子系统间互连。根据网络连接的子系统不同，网络系统分为计算网络、存储网络、集群管理和监控网络。



计算网络



存储网络



集群管理网络



集群监控网络



KVM系统

基础架构：为整个集群系统提供供电、散热、布线等服务。基础架构包括机柜、电源系统、散热系统和布线系统，聚合整个集群系统的基础设施同时为集群提供供电、散热和布线等基础服务。



机柜



电源系统



散热系统



布线系统

英特尔®至强®可扩展平台
加速探索与获取洞察
咨询电话 400 819 6776



软件系统：用于整合管理硬件资源，为用户提供软件平台与服务。联想高性能集群系统提供丰富的软件系统层来支撑集群硬件系统向用户提供服务。联想高性能计算软件系统包括快速集群部署、易于集群管理、合理资源调度、实时监控集群状态等多种工具，同时又集成很多不同领域热门用户应用及相应的开发环境。

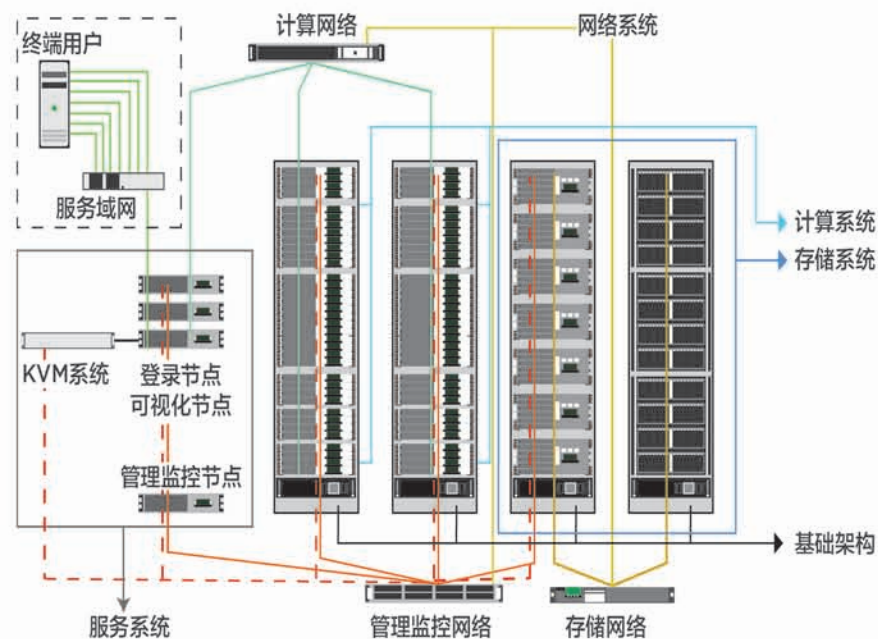


图 联想高性能计算高性能集群硬件系统架构图

联想高性能集群系统构成						
子系统	子系统功能模块					
计算系统	CPU 计算节点	GPU 计算节点	大内存节点	N/A	N/A	N/A
存储系统	I/O 节点 + 盘阵	DF 存储系统	DSS-G 并行存储	N/A	N/A	N/A
服务系统	管理节点	监控节点	登录节点	可视化节点	KVM 系统	N/A
网络系统	计算网络	存储网络	管理网络	监控网络	N/A	N/A
基础架构	机柜	电源系统	散热系统	布线系统	N/A	N/A
软件系统	集群软件系统	节点操作系统	编程环境	应用软件	文件系统	测试工具

表 联想高性能计算集群系统构成表



图 联想高性能集群系统的软件系统层次图

- 节点操作系统：集群中各子系统节点支持多种操作系统，如 Red Hat Enterprise Linux、CentOS、SUSE、Windows Server 等。
- 集群软件系统：集群软件系统提供集群的管理、集群的监控、用户作业和集群资源的调度、集群的优化调试等功能。
- 文件系统：存储系统的软件系统，用于存储资源的整合管理和满足计算系统的存储需求。
- 编程环境：用于支撑用户计算程序的编写和运行，以及集群调试优化和测试程序的运行。
- 用户应用软件（应用软件）：用户生产工具。

常见应用特性分析

联想高性能计算系统集成和支持了多个领域的多种商业和开源的科学计算应用软件，包括商业应用软件如 MOL-PRO、GAUSSIAN、Q-CHEM、MATLAB、VASP、TURBOMOLE、WIEN2K、SPARTAN、Materials Studio、Linda、LS-DYNA、FLUENT、CRYSTAL、ADF 等；开源应用软件如 GROMACS、GrADS、ABINIT、GMT、CASTEP、CPMD、DL_POLY、DOCK、GRAPES、MM5、NAMD、AUTODOCK、VMD、TINKER、Open-MX、VENUS、WRF 等。

英特尔® 至强® 可扩展平台
加速探索与获取洞察
咨询电话 400 819 6776



(1) MOLPRO

MOLPRO 是国际上广泛使用的专业级电子结构量子化学计算软件，MOLPRO 的特色是高精度计算，通过多参考 CI，耦合簇和有关的方法，广泛处理电子相关问题。直接积分局域电子相关方法使得同分子尺寸正相关的计算量大为减少，从而能够对更大的分子体系进行准确的从头计算。MOLPRO 基于 Linux 操作系统运行，程序主体是用 Fortran 90 语言编写，支持大多数常见的 MPP 超级计算机和计算集群。对于分布式内存系统如集群，由 GlobalArrays 并行工具集支持并行运行；对于共享内存系统，由 OpenMP 的方式实现了部分代码的并行化。需要注意的是，MOLPRO 的并行部分通讯代价较大，一般不建议基于普通以太网运行。

MOLPRO 中并行化了传统积分的 SCF、DFT、MRCI、MP2、LMP2、CCSD(T) 能量和 SCF、DFT 梯度计算，以及直接积分和密度拟合的 SCF、DFT、LMP2 和 LCCSD(T) 计算。在最新版本中还并行化了 DF-HF、DF-KS、DF-LCCSD(T)。

MOLPRO 目前的并行代码包括两种形态，一是运行于分布式内存的多处理器系统上，如计算集群，这种方式需要 GlobalArrays 并行工具的支持；另一种方式是运行在共享内存系统上，这部分代码采用 OpenMP 进行并行化。由于计算模型导致的并行任务间通讯负载较大的问题，MOLPRO 并行版在普通网络上难以表现出良好的加速比，通常建议其在专用高速网上运行。

MOLPRO 通常有三种编译模式：

串行模式：运行时只有单个实际执行体；

MPP 模式：一个任务被分为多个执行实体，同时运行在多个处理器上；

MPPX 模式：多个任务分配到多个处理器上同时执行，每个任务用一个处理器串行执行，例如，在梯度的有限差分求解中，多个构型计算在多个处理器上同时进行，处理不同的数据，属于 SPMD 的计算模式。就目前而言，这种模式只用于数值求一阶梯度和 Hessian。

MOLPRO 的运行形态在编译时指定为上述三种模式之一，并且会在输出结果中显示。

(2) GAUSSIAN

Gaussian 是目前最流行的量子化学综合软件包，用于研究许多化学领域的课题，例如取代基的影响，化学反应机理，势能曲面和激发能等。

Gaussian 用 Fortran 语言编写，从量子力学的基本原理出发，可计算能量、分子结构、分子体系的振动频率以及大量从这些基本计算方法中导出的分子性质。它能用于研究不同条件下的分子和反应，包括稳定的粒子和实验上难以观测的化合物，例如瞬时的反应中间物和过渡结构。

Gaussian 03 软件主要采用两种并行方式，对于 IBM 小型机或者 SGI 的 Origin 系列或者 Altix 系列的高性能机来说，采用基于 OpenMP 的并行方式。OpenMP 模式下的多个线程并发地运行在基于共享内存方式或者基于 NUMA 体系结构的高性能计算机上，以此提升程序的运算速度和运算规模。这种并行方式受到单节点的处理器数的限制。如果服务器的处理器数为 4，则 OpenMP 并行方式最多只能使 4 颗处理器并行工作。

对于分布式内存的集群系统，Gaussian 03 用 Linda 并行库来实现跨节点的并发执行。TCP-Linda 是专门为 Gaussian 03 设计，用以实现其分布式内存系统下的并行。Gaussian 03 在提交任务的时候选择分布式内存方式还是共享内存方式。

Linda 模型定义了一个功能强大的逻辑存储器 (TS) 和在其上的一组核心操作 (in, out, read, eval)，它们能够方便地嵌入到不同的语言 (如，C, Fortran) 中而构成相应的并行语言 (C-Linda 和 Fortran-Linda 等)，该模型同时还支持动态程序设计和两种不同的编程风格 (Master/Slaver 和 Divide-and-Conquer)，为用户开发不同类型的应用程序提供了灵活的手段。

TS 是一个可被多个进程共享的、能同时存放数据和任务的数据箱。进程从 TS 中提取任务或数据进行计算，并将结果或生成的新任务放入 TS 中。并行执行的进程之间通过 TS 进行间接的通信和同步。

(3) Q-CHEM

Q-CHEM 是由遍及全球的多个研究机构共同开发的从头计算量子化学软件包，它包含了当前流行的各种最先进量子化学理论方法和工具。Q-CHEM 的特点是与当前最先进的量子化学理论和计算方法结合的非常紧密。Q-CHEM 采用了全新的方法理论、最先进的算法和现代编程技术进行设计，系统地组合各种计算方法和工具，极大地提高了计算速度和准确性，能够以较高精度进行大分子体系的运算。

Q-CHEM 主要用作电子结构从头计算，可以对分子的基态和激发态进行第一定律计算。Q-CHEM 可以采用串行和并行两种方式运行。Q-CHEM 的 HF、DFT 方法及它们的二阶导数计算。采用 MPI 编写，使用分布式模式并行运行。Q-CHEM 对解析和数值积分的计算采用的动态负载均衡的办法，并且在求解耦合扰动 SCF 方程中采用全局内存使用共享的方式，使得它能够利用并行计算机的分布内存进行大结构的频率计算。

(4) MATLAB

MATLAB 是建立在向量、数组和矩阵基础上的一种分析和仿真工具软件包，自 1984 年由美国 MathWorks 公司推向市场以来，历经二十多年的发展与竞争，已成为国际公认的最优秀的工程应用开发环境。

MATLAB 包含各种能够进行常规运算的“工具箱”，如常用的矩阵代数运算、数组运算、方程求根、优化计算及函数求导积分符号运算等；同时还提供了编程计算的编程特性，通过编程可以解决一些复杂的工程问题；也可绘制二维、三维图形，输出结果可视化。

MATLAB 在 2004 年拥有了第一代的并行与分布式处理产品分布式工具箱 (Distributed Computing Tools)。分布式工具箱将待处理任务分配给多个处理器，使采用不同运算结构的 MATLAB 用户共享同一个处理器池的计算能力。在最新的版本中，MATLAB 能使跨节点运行的任务之间相互通信，做到真正的并行运算，此功能是通过 MPI 实现。分布式计算工具箱可以在多处理器计算环境中使用 MATLAB 解决计算、数据密集型问题。

(5) VASP

VASP 是固体物理程序中非常稳定，性能优秀的软件之一，被广泛的用于固体性质的计算，一直是相关领域研究人员的关注热点。VASP 使用赝势和平面波基组，进行量子力学分子动力学计算，它基于 CASTEP 1989 版，由奥地利 Vienna 大学开发。

VASP 中采用基于有限温度下的局域密度近似 (用自由能作为变量) 的方法，对每一 MD 步骤用有效矩阵对角方案 and 有效 Pulay 混合求解瞬时电子基态。这些技术可以避免原始的 Car-Parrinello 方法 (基于电子、离子运动方程同时积分的方法) 存在的一切问题。离子和电子的相互作用采用超缓 Vanderbilt 赝势 (US-PP) 或投影扩充波 (PAW) 方法描述，可以在相当程度上减少过渡金属或第一行元素的每个原子所必需的平面波数量。同时，VASP 可以很容易地进行力与张量的计算，以达到把原子衰减到其瞬时基态的目的。VASP 的并行由 MPI 实现，可以在分布式内存环境上进行并行计算。

(6) TURBOMOLE

TURBOMOLE 是一个强大的量子力学软件包，由德国卡尔斯鲁厄 (Kalsruhe) 大学的 Ahlrichs 教授领导的理论化学研究小组开发而成。TURBOMOLE 经过了十多年的不断改进，已经成为化学、物理和工程人员的有力工具。

与众多同类软件不同，TURBOMOLE 着重解决如何用尽量短的时间和尽量少的内存需求，快速稳定地处理工业应用型的分子。特别是独有的 RI-DFT 方法，可以较其它大多数量子力学程序节省 10 倍的 CPU 时间。由于它使用了辅助函数 (单元元分解近似使得 TURBOMOLE 非常适合于计算大分子或重复计算中等分子)，TURBOMOLE 可以以串行和并行两种方式运行。并行代码用 MPI 编写。

(7) WIEN2K

WIEN2K 采用密度泛函理论来计算固体的电子结构，是基于键结构计算最准确的方案——完全势能 (线性) 增广平面波 ((L)APW)+ 局域轨道 (lo) 方法。在密度泛函中可以使用局域 (自旋) 密度近似 (LDA) 或广义梯度近似 (GGA)。

WIEN2K 主要运行在 Unix/Linux 上，有串行和并行方式。

WIEN2K 程序的 I/O 操作较频繁，可能导致处理器的闲置状况，通常解决方案是扩充读写缓冲，减少对磁盘的访问。WIEN2K 中某些代码对处理器资源的消耗极大，其中一些已经用 MPI 方式加以并行化。WIEN2K 的并行模式分 K 点 (K point) 并行和细粒度 (fine grained) 并行，二者属于不同层次上的并行划分，在安装时的并行机选项中可以指定具体的并行方式。

(8) SPARTAN

SPARTAN 是由美国 Wavefunction 公司研发的多功能化学软件，是化学领域强大的分子力学和量子力学计算程序。SPARTAN 具有优秀的图形显示功能和成熟的算法，为制药与生物技术机构在靶标识别与确立、先导物的选择与优化和工艺优化过程提供关键数据支持。

SPARTAN 采用的理论方法有从头 Hartree-Fock、密度泛函 (局域密度近似，以及 BP, BLYP, EDF1, B3LYP)、MPn (MP2, MP3, MP4, LMP2)、精细的相关能计算 (CCSD, CCSD(T), OD, OD(T), G2, G3)、半经验方法 (MNDO, MNDO(d), AM1, PM3)、蒙特卡罗和分子力学。

SPARTAN 具有强大的图形和用户界面功能：从计算的 IR 数据绘制红外光谱图、从 Hartree-Fock 分子轨道理论绘制 ¹³C NMR 图、分子数据库及搜索、氢键显示、内反应坐标 (IRC) 的产生和追踪、SDF, TGF, MOL 各式的分子坐标文件的输入输出、动画速度的人为控制和键型显示等。

(9) Materials Studio

Materials Studio 是 Accelrys 公司专为材料科学领域开发的新一代材料计算软件，用以帮助研究人员解决化学及材料工业中的许多重要问题。Materials Studio 软件采用 Client/Server 结构。Materials Studio 将高质量的材料模拟带入了个人电脑的时代，极大的降低了研究的门槛。

多种先进算法的采用使得 Materials Studio 在性质预测、聚合物建模和 X 射线衍射模拟等方面具有强大的功能。同时，Materials Studio 可以生成的高质量图片、能处理各种来源的图形、文本以及数据表格。

Materials Studio 采用客户端 / 服务端的工作模式，较小计算量的任务直接运行在客户端，而长时间或大运算量的计算则运行到服务器端。

Materials Studio 中许多模块都实现了并行化，其中 Dmol3 模块是世界上最快的从头计算代码之一，也已成为制药业的一个重要的工具。而 CASTEP 模块的并行版本则可以模拟包含数百原子的大系统。

Materials Studio 的所有模块都是通过第三方队列软件支持串行工作提交，比如主流的 OpenPbs 和 LSF。对于 Linux 平台，还支持跨节点的 MPI 并行方式。

(10) Linda

Linda 是美国耶鲁大学与科学计算协会共同研制的用于实现并行程序设计的、与机器无关的程序环境。Linda 通过增加一些函数对传统的程序设计语言进行扩充（如 C、Fortran 等），使其能实现并行程序的设计。将 Linda 映射到各计算语言中，就形成了可进行并行计算的程序语言 C-Linda、Fortran-Linda。Linda 可以运行在共享和分布式内存系统上。

Linda 运行在分布存储多处理机上，引入了虚拟共享存储概念，通过在各处理机上实现一个虚拟的共享存储器，将分布存储多处理机模拟成为基于共享存储器的多处理机，从而支持共享存储编程方式，提供比消息传递方式更方便的并行编程模式。

Linda 的虚拟共享存储器称为元组空间（Tuple Space）。元组空间由一组有序的元组（tuple）组成，元组的每个域都包含有实际的数据。元组空间是相联存储器，元组的标识与选择是通过域值匹配，而不是通常采用的地址选择方法。

Linda 提供了以下 4 种对元组的基本操作：

- out: 将数据放入元组空间，整个操作是顺序进行的；
- eval: 功能同 out，但它是并行执行的；
- in: 从元组空间中选择匹配的数据，并将数据从元组空间中删除；
- rd: 功能同 in，但它不将数据从元组空间中删除。

当两个进程需要交换数据时，不是采用消息传递来直接通信，而是通过读写元组空间来完成（即共享内存系统的并行程序设计）。例如，如果 A 进程需要向 B 进程传递数据，则需要完成以下的操作：A 向 Tuple 空间写一个数据，B 在 Tuple 空间中读取需要的数据。Linda 在编译系统和运行系统上做了很多优化工作，使用 Linda 编写的应用程序在效率上能够接近用传统消息传递方式编写的应用程序。

TCP-Linda 是专门为 Gaussian 03 设计，实现 Gaussian 03 在分布式内存系统上并行的必须程序，使 Gaussian 03 能突破单节点的限制，充分利用可用的计算资源，对于极耗资源的量子化学计算来说，具有相当的实用意义。

(11) LS-DYNA

LS-DYNA 最初是由 John O. Hallquist 博士在美国劳伦斯·利弗摩尔国家实验室开发。LS-DYNA 最早实践了显式有限元理论，当代所有的显式求解程序都源于 LS-DYNA。

LS-DYNA 以拉格朗日（Lagrange）算法为主，同时有欧拉（Euler）求解技术；以显式求解为主，兼有隐式求解功能；以结构分析为主，兼有热分析、流体 - 结构耦合功能；以非线性动力分析为主，兼有静力分析功能（如动力分析前的预应力计算和薄板冲压成形后的回弹计算）；特别适合求解各种结构的高速碰撞、爆炸和金属成型等高度非线性瞬态动力学问题。

LS-DYNA 主要应用在如下领域:

- 汽车工业: 碰撞分析、乘员约束系统及气囊设计、乘客被动安全、部件加工、轮胎在积水路面排水性和动平衡分析;
- 国防与军工: 内弹道及终点弹道效应、装甲和反装甲系统、穿甲弹与破甲弹设计、战斗部结构设计及起爆参数 / 杀伤元素优化、冲击波传播、空气 / 水 / 土壤与容器中爆炸、爆炸容器的设计优化分析、激光束 / 定向能等多种动态载荷加载模拟分析、爆炸对建筑物等设施结构的破坏分析等;
- 电子工业: 跌落冲击分析、包装设计、热分析、电子封装、电子产品抗冲击性设计;
- 航空航天: 鸟撞、叶片包容性设计、异物损伤分析、飞机结构冲击动力分析、碰撞、坠毁、冲击爆炸及动态载荷、特种复合材料设计、制造工艺仿真及优化;
- 建筑业: 地震安全、混凝土结构、爆破拆除、公路桥梁设计;
- 石油工业: 液体晃动、完井射孔、管道设计、爆炸切割、事故模拟、海上平台设计。

LS-DYNA 具有良好的并行设计, 能够支持分布式及共享内存方式的并行。LS-DYNA 能够支持网格计算 (Grid Computing), 将多台计算节点组合成多个网格点以实现协同计算。

在用户模型定义准确、单元规模和处理器数目匹配的情况下, LS-DYNA 的性能加速比和处理器数目基本呈线性关系。

LS-DYNA 得到了主流作业调度软件 Platform LSF 的有力支持。在共享内存和分布式内存系统上, LS-DYNA 作业都具有检查点 (checkpoint) 和重启功能。

(12) FLUENT

FLUENT 软件最初原发于英国谢菲尔德大学 (Sheffield University), 现已成为一款商用 CFD 软件。FLUENT 占有全球商用 CFD 软件领域第一的市场份额, 在航空航天、旋转机械、航海、石油化工、汽车、能源、计算机 / 电子、材料、冶金、生物、医药等领域有广泛的应用, 美国宇航局 (NASA)、美国国防部 (DOD)、美国能源部 (DOE)、ABB 公司、西屋公司、波音公司、福特公司、三菱公司、IBM 公司、杜邦公司及各大汽车厂商等都是 FLUENT 的用户。

FLUENT 主要用于计算流体流动和传热问题。其提供的非结构网格生成程序, 对相对复杂的几何结构网格生成非常有效。FLUENT 具有网格自适应能力, 能高效地精确求解大梯度的流场。

FLUENT 采用的是客户端 / 服务器 (Client/Server) 设计模式。用户在客户端进行模拟问题的设定和数据输入, 然后将计算任务提交服务器; 服务器进行模拟计算并将结果传回客户端。

FLUENT 的并行计算主要是在服务器端进行的。FLUENT 并行运算由一个主机进程和多个节点进程组成。节点进程的数量和系统的处理器数量有关, 计算任务均由节点进程完成。FLUENT 能支持共享式和分布式内存系统上的并行计算。

在网格处理技术方面, FLUENT 在并行环境下可以实现网格的自动分区并动态地平衡计算负载。

FLUENT 的并行效率和互连网络有很大关系。核心的线性方程组求解过程产生频繁的信息传递和交换, 因而使用如 InfiniBand 等高速网能有效地提高程序的扩展性。ICSGI、HP 和 Intel 等多种优化 MPI 接口也对 FLUENT 的并行性能有非常正面的影响。此外, FLUENT 中不同算法在共享式或者分布式系统上也会表现出一定的性能落差。

(13) CRYSTAL

CRYSTAL 是研究晶态固体的最流程序之一, 并且是第一个公开发表的程序。CRYSTAL 程序密度泛函或各种混合近似方法, 计算周期体系的电子结构。周期体系的 Bloch 函数作为原子中心 Gaussian 函数的线性组合展开。采用强大的屏蔽技术研究实空间区域。

CRYSTAL 可以用于研究分子, 多聚物, 表面及晶体的物理和电磁性质, 使用全电子基组或者价基组进行限制性闭壳层、限制性开壳层或者非限制性计算。可以自动操作空间对称性并对分子提供点群对称性和平移对称性。

CRYSTAL 具有如下功能:

<p>01</p> <p>哈密顿量计算, 求解基于数值求积分方案的数值网格并进行基于数值方案的密度拟合</p>	<p>02</p> <p>能量导数计算</p>	<p>03</p> <p>提供单点能计算和自动几何优化等计算类型</p>	<p>04</p> <p>支持多种基组: Gaussian型基组、用户自定义赝势基组、Hay-Wadt大核和小核基组等</p>
<p>05</p> <p>独有的周期体系;</p>	<p>06</p> <p>多种波函数特性与分析: 能带结构、态密度、全电子电荷密度、电场梯度、结构因子、静电势及其导数等</p>	<p>07</p> <p>CRYSTAL已经实现了代码的并行化, 以MPI实现, 能够在分布式内存系统上进行并行计算。</p>	

英特尔®至强®可扩展平台
加速探索与获取洞察
咨询电话 400 819 6776



(14) ADF

ADF 是一个通用密度泛函程序包，基于密度泛函理论 (DFT)，主要应用于量子化学计算。广泛应用于医药化学、材料学等研究及应用领域，尤其是同类和异类催化、无机化学、重元素化学、生物化学及多种光谱学。

ADF 的优势主要体现在计算速度和对过渡金属的处理上面。主要具有以下功能：几何优化、过渡态、反应路径及红外频率的表征及计算；分子性能测试的分析与表征；模拟溶剂、蛋白质及其它物化环境；强大的图形用户界面；进行聚合物、异类催化剂及大量结晶的研究。

ADF 主要有如下特色：

- 精度方面：精确的、可调节的数值积分方案；应用于整个周期表的，最高到四 zeta 的 Slater 型轨道；无需赝势或 ECP 近似；对过渡金属化合物稳定的 SCF 收敛性；现代的交流相关能泛函和势能泛函；
- 独有的 slater 轨道：采用 Slater 型基组 (STO)；正确反映原子轨道的尖峰区和渐进行为；BAND 采用数值的轨道和 STO 轨道。

ADF 以单程序多数据 (SPMD) 的模式实现并行计算，在处理器中分配格点和原子对，然后对其执行相同的计算。ADF 的并行用消息传递机制实现，基于 MPI 和 PVM 开发了专用的并行库，极大地提高了通讯的效率。通常情况下，高达 90% 以上的 CPU 时间都被用来进行数值积分。

(15) GROMACS

GROMACS (<http://www.gromacs.org/download/index.php>) 是由 Groningen 大学开发一个分子动力学模拟包，在基于 Fortran77 的 GROMOS 的基础上，加入了 C 语言和大量的优化代码。GROMACS 支持所有的当代分子动力学应用算法，其一个著名的用途是进行快速准确的蛋白质模拟。

GROMACS 具有如下特点：

- 性能远高于同类软件：将最内层循环的针对不同体系结构进行优化；
- 交互界面友好：提供清晰的拓扑和参数文件；进行及时的进度提示；能用 gzip 读取压缩文件；提供简单实用的命令行控制方式；
- 数据文件独立于硬件体系，无须区分大小端 (big/little endian)；版本间数据文件格式兼容；
- 能够以并行方式运行，并行代码用 MPI 编写而成；
- 精巧的算法保证对性能调节不会影响计算过程和结果的准确性；

- 能够自动的构建蛋白质分子拓扑；
- 正在向量子化学和生物信息学进行扩展。

GROMACS 代码中对部分算法进行了并行化。例如对 MD (Molecular Dynamics) 算法，对数据进行一维的分割，各计算点只和相邻点发生数据通讯，这样可以使并行代码编写简单化，而且还能将琐碎的通讯合成大数据量的通讯，由此减少通讯的额外开销。除 MD 算法外，GROMACS 还对 FFT 和排序算法进行了并行化。

(16) GrADS

GrADS (<http://www.iges.org/grads/>) 是用于气象数据处理和显示的交互式工具，通过其提供的集成环境可以容易地对数据进行读取、加工、图形化显示和打印输出，同时具有分析功能强，坐标丰富，显示速度快等优点，因此已经成为国内外气象数据显示方面通用的标准图形环境之一。

GrADS 既可以处理站点数据，也可以处理格点数据。GrADS 对数据的处理使用四维数据环境，包括了三维空间和时间。在 GrADS 中，原始数据和元数据分别存储在不同的文件中。原始数据文件保持的是纯数据，没有空间和时间标识。元数据文件称为数据描述文件，包含了对原始数据文件的描述以及如何获取等信息。所有数据集通过数据描述文件被放置到四维空间中。GrADS 可以识别二进制的格式，也可以采用 GRIB, NetCDF 或 HDF-SDS 等通用的数据交换格式。

用户可以在交换环境下，通过在命令行输入类似 Fortran 语句的表达式进行操作。可以将键入的命令记录到一个文本文件中，通过 exec 命令批处理执行。同时 GrADS 提供了一种脚本语言作为编程界面，解释器就是 GrADS 本身，由 run 命令控制执行。该语言提供了变量、流控制、输入输出等高级语言功能，用户可以容易地以编程方式处理复杂的分析和图形显示。除了能够批量方式运行 GrADS 处理冗长的数据外，使用脚本还可以让用户显示按钮或下拉菜单，并根据用户选择执行相应的动作。

GrADS 提供了丰富的内建函数，用户也可以借助任何编程语言添加自己的函数。GrADS 支持多种图形化显示方式，并可以 postscript 或其他图形格式输出。

GrADS 目前版本为 1.8s11 和 1.9b4，支持包括 Linux, Mac, IRIX, AIX, Windows 在内的大多数主流操作系统。

(17) ABINIT

ABINIT (<http://www.abinit.org/>) 是由 Xavier Gonze 所领导、世界上众多电子结构计算研究人员集体开发的平面赝势法电子结构计算程序包。

程序包基于密度泛函理论 (Density Functional Theory, DFT), 采用赝势和平面波基矢的方法来处理由电子和核所组成的体系, 它可以计算体系的总能、电荷密度以及电子结构。

ABINIT 也可以按 H-F 力和压力来优化体系的几何结构, 或进行根据这些力进行分子动力学模拟或计算得到动力学矩阵、Born 有效电荷及介电张量。在密度泛函理论的框架下可以计算分子体系的激发态, 或基于多体微扰论 (GW 近似) 来处理激发态。除了主要计算模块外, 程序包也提供了一些不同的工具模块用来处理计算结果和数据。

ABINIT 具有如下计算功能:

- 通常的电子结构计算 (晶格参数、结合能、体弹性模量、能带结构、电子态密度和电荷密度分布);
- 分子动力学模拟;
- 晶体磁性和磁结构的计算;
- 基于线性响应或密度泛函微扰理论计算声子色散曲线、声子态密度和热力学量, 介电性质、Bohr 有效电荷、压电系数和弹性常数、电 - 声耦合常数;
- 计算线性和非线性的光学性质;
- 采用 GW 近似对半导体的 LDA 带隙进行修正;
- 基于含时密度泛函理论计算体系的激发态的性质。

(18) GMT

GMT (<http://www.soest.hawaii.edu/gmt/>) 是大约 60 多个 UNIX/LINUX 工具的集合。它使得用户可以处理二维及三维的数据集 (包括 filtering, trend fitting, gridding, 投影等)。GMT 支持 30 多种常用的投影及线性、对数和指数拉伸, 并且支持诸如海岸线、河流及行政界线等。

GMT 系统最早由哥伦比亚大学的 Wessel 和 Smith 开发, 全球大约有 6000 多位的使用者。早期是在 UNIX 系统上开发, 可以和其它许多 Unix 的工具一起使用。GMT 主要是在指令模式下执行。GMT 是免费软件, 数据输出的主要格式是 PS (postscript file), 因此几乎在各种平台上都可支持。

(19) BLAST

BLAST (<http://www.ncbi.nlm.nih.gov/BLAST/>) 是基于局部序列排比的常用数据库搜索工具, 建立在严格的统计学的基础之上的, 它集中于发现具有较高的相似性的局部比对。

BLAST 具有如下特点:

使用方便、功能齐全

速度快、结果可信度高

配套数据库不断更新

提供web上的
免费服务

BLAST算法本身具有
良好的并行性, 其并
行版基于MPI

(20) CASTEP

CASTEP (<http://www.tcm.phy.cam.ac.uk/castep/index.html>) 是由剑桥大学凝聚态理论研究组开发的一套先进的量子力学程序, 可以进行化学和材料科学方面的研究。CASTEP 一个基于密度泛函方法的从头算量子力学程序, 可以模拟固体、界面和表面的性质, 适用于多种材料体系, 包括陶瓷、半导体和金属等。

基于总能量赝势方法, CASTEP 根据系统中原子的类型和数目来进行各种性质的预测。CASTEP 使用的平面波赝势技术已经过实践的考验, 每年发表的数百篇科学文献充分说明了其在许多领域中的成功应用。

CASTEP 主要具有如下功能:

- 基于总能量赝势方法, 预测出包括晶格常数、几何密度、弹性常数、能带、态密度、电荷密度、波函数以及光学性质在内的各种性质;
- 电子的交换和相关效应采用局域密度近似 (LDA) 和广义密度近似 (GGA);
- 电子波函数用平面波基组扩展 (基组数由 Ecut-off 确定);
- 分子轨道波函数采用原子轨道的线性组合 (LCAO) 构成。

(21) CPMD

CPMD (<http://www.cpmd.org/>) 是以 CPMD (Car-Parrinello Molecular Dynamics) 方法为基本思想所编写, 是 1993 年 IBM 瑞士研究所的一个计算物理研究计划的产物。

CPMD 是一种量子化学计算方法, 将分子动力学和密度泛函有机结合, 最大限度的发挥了密度泛函的优势。CPMD 是从头计算法的延伸, 在精确模拟和简化模型上有独特的优势。

CPMD 主要应用于固体和液体, 静力学和动力学, 固态物理和固态化学。

CPMD 主要有如下特点和功能:

- 孤立体系和周期边界体系的计算;
- 对分子和晶体使用对称性;
- 波函数优化;
- 几何优化与过渡态寻找;
- 恒定能量, 恒温, 恒压的分子动力学;
- 路径积分分子动力学;
- 线性响应函数计算, NMR, Raman 和 IR;
- TDDFT 计算激发态;
- 多电子特性;
- 周期体系的偶极矩;
- 激发态的分子动力学;
- QM/MM 方法;
- OpenMP 并行化。

(22) DL_POLY

DL_POLY (http://www.cse.scitech.ac.uk/ccg/software/DL_POLY/) 是串行和并行分子动力学模拟软件包, 由 Daresbury 实验室的 W.Smith 和 T.R.Forester 开发。

DL_POLY 可以串行和并行两种方式运行。对于并行方式, 既能以 MPP 模式并行, 也能在分布式内存系统上运行。

DL_POLY 可以有两种并行的方式:

- 复制数据的方法, 适合在数十量级的处理器上模拟十万以下个数的原子;
- 区域分解的方法, 能够在 8 至 1024 个处理器上, 模拟百万量级个数的原子。

对于一个 DL_POLY 许可, 同时提供两个版本。DL_POLY 还提供基于 JAVA 语言的图形用户界面。

(23) DOCK

DOCK (<http://dock.compbio.ucsf.edu/>) 是 Kuntz 研究小组发展的分子对接程序, 可能是目前应用最为广泛的分子对接程序之一。它能自动地模拟配体分子在受体活性位点的作用情况, 并把理论预测最佳的方式记录下来。而且该方法能够对配体的三维结构数据库进行自动搜索, 因此被广泛应用于基于受体结构的数据库搜索的药物设计中。

用 DOCK 进行药物设计以及数据库的搜索基本上可以分为下面几个步骤:

配体和受体相互作用位点的确定, 评分系统的生成, DOCK 计算及 DOCK 结果的处理与分析。活性位点的确定和表达是 DOCK 最重要的特点之一。活性位点特征的确定对于 DOCK 研究是非常重要的, 因为配体分子和受体相互作用过程的模拟主要就是参考几何位点的几何特征进行的。在 DOCK 中, 活性位点的确定通过 sphgen 程序来完成。DOCK 软件包中 sphgen 程序生成受体表面所有的凹陷的负像, 并对这些负像进行聚类分析。在 DOCK 程序中, 表面点采用了 Richards 提出的模型。在这些表面点的基础上, 采用 sphgen 程序生成了负像, 它实际上由一些与分子表面点相切的圆球叠加而成。

在生成负像的基础上, 就可以进行配体分子和活性口袋之间的匹配。在这里, 配体也采用一组球集来表示, 和负像不同的是, 配体所用的球集表示配体所占的空间区域。如果配体分子能和活性口袋形成比较好的匹配, 那么配体的球集一定能和活性口袋中的负像形成好的叠合。配体分子和负像之间的匹配原则是基于配体和受体之间球集的内坐标的比较。

按照匹配原则得到了配体和受体之间的匹配情况之后, 就要通过合理的得分函数来选择最优的结果。DOCK 提供了多种得分函数来评价配体和受体之间的结合情况, 包括原子接触得分以及能量得分。

DOCK 进行分子对接时, 配体分子可以是柔性的。对于柔性的分子, 其键长和键角保持不变, 但可旋转二面角是可以发生变化的。在 DOCK 中, 柔性分子的构象变化通过下面的操作实现: 首先是刚性片断的确定, 然后是构象搜索。构象搜索采用两种方法: 一种是锚优先搜索 (anchor-first search), 第二种方法是同时搜索 (simultaneous search)。

(24) GRAPES

GRAPES 全球区域同化预报系统是中国气象局负责研究开发的、研究与业务通用的数值天气预报系统，属于“十五”国家科技攻关计划。

GRAPES 是集常规与非常规变分同化、静力平衡与非静力平衡、全球与区域模式、科研与业务应用、串行与并行计算、标准化与模块化程序、理想实验与实际预报等为一体，中小尺度与大尺度通用的先进数值预报系统。

GRAPES 作为我国自主研发的天气预报模型，具有下列创新点：

- 可以把空中和地面的资料全部用到数值预报中，改变了国家花费几十亿元发展的气象卫星和多普勒天气雷达闲置的情况；
- 性能优于已有的国外模式，对于暴雨、强对流天气等可以有更好的预报结果；
- 较高的预报准确率，在可用预报时效和预报精细程度方面有很好的表现；
- 具有完全自主知识产权，改变了过去系统受制于国外的局面。

GRAPES 应用非常广泛，目前国内很多气象部门已经在对 GRAPES 进行研究和应用。

(25) MM5

MM5(<http://www.mmm.ucar.edu/mm5/>)是美国宾州州立大学和美国国家大气研究中心(NCAR)在 20 世纪 90 年代初研制发展的新一代中尺度天气模型。中尺度强对流天气系统尺度小，具有突发性和多发性特点，且天气变化激烈，造成的灾害范围广。MM5 的出现对于区域中尺度数值天气预报的研究有着重大的意义。

MM5 是建立在多重嵌套网格上的格点差分模式。MM5 模式是一种双向嵌套有限差分预报模式，其计算域按照模式预报的不同要求和分辨率分成母域和嵌套域，嵌套域层数可以达到 9 层。MM5 模式计算涉及到粗网格（母域）和细网格（嵌套域）两个计算区域，数据关系复杂。MM5 有限区域差分模式的所有物理量的计算都是在网格空间上进行的。每一个物理量的模式积分在每个水平经纬网格点是相互耦合的，但在垂直层上无数据相关性。

大气运动的物理特征非常适合于并行处理，如大气中的水平和垂直运动、各种非绝热物理过程等，在一定程度上都可以看作相互独立的。

MM5 基于消息传递机制实现并行计算，低层依赖 RSL (RuntimeSystemLibrary) 库实现。RSL 库是基于消息传递标准 MPI 之上的并行计算应用程序接口，主要针对有限差分格式网格计算。

MM5 采用水平二维数据划分实现分布式并行计算。算法的基本思想是：（1）将处理机划分为 $N_x \times N_y$ 的逻辑网格单元，分别对母域和嵌套域的计算网格点进行经向（x 方向）和纬向（y 方向）二维剖分，将各个网格层上的计算划分为 $N_x \times N_y$ 个逻辑计算单元，分布到各处理机上进行；（2）采用 RSL 库特殊的数据结构，计算出每个逻辑计算单元之间存在数据相关性的网格点，实现网格层内部及粗 / 细网格层之间的通讯。

MM5 的数据通讯包括域内通讯和域间通讯两种。域内通讯是指计算域（母域或嵌套域）内部由于数据划分，使计算分布在各个处理结点上进行，相邻计算子域之间需要调用 RSL 库显式地通过消息传递来进行数据的通讯。域间通讯是母域和嵌套域之间由于“动力”和“反馈”的相互作用而进行的数据交换。

MM5 采用一种基于不规则区域的数据划分策略，使用该方法能够基本保证负载均衡。理想情况下，采用不规则区域数据划分方法，基本能将网格点均匀的分配到各个处理器上。

(26) NAMD

NAMD (<http://www.ks.uiuc.edu/Research/namd/>) 是一款并行的分子动力学计算软件, 曾获得 2002 年度的金铃奖 (Gordon Bell Award), 用于在大规模并行计算机上快速模拟大分子体系。NAMd 用经验力场, 如 Amber, CHARMM 和 Dreiding, 通过数值求解运动方程计算原子轨迹。用于预测生物分子的动力学行为和重要性质, 如弥散因子和内聚能等等。

NAMD 具有如下基本功能:

- 以两种任务类型提交运算: 几何优化和分子动力学;
- 有效共轭梯度最小化;
- 周期和非周期分子动力学模拟;
- 速率重新标度、球边界条件和谐振原子限制;
- 恒温 and 恒压的分子动力学模拟、热平衡分子动力学模拟;
- 多时间步计算;
- 化学和构象自由能计算;
- 基于 Tcl 的脚本;
- 两种计算特性: 均方位移和内能分解。

NAMD 可以以串行和并行两种方式运行。NAMd 的并行机制基于 Charm++ 对象, 可以在高端的并行机上进行上千个处理器以上规模的并行计算。对于使用普通以太网的分布式内存系统, 可以进行数百个处理器规模的运算。

(27) WRF

WRF (<http://www.wrf-model.org/index.php>) 模式系统是由许多美国研究部门及大学的科学家共同参与进行开发研究的新一代中尺度预报模式和同化系统。

WRF 模式是一个完全可压非静力模式, 控制方程组都写为通量形式。网格形式采用 Arakawa C 格点, 有利于在高分辨率模拟中提高准确性。模式的动力框架有三个不同的方案。前两个方案都采用时间分裂显式方案来解动力方程组, 即模式中垂直高频波的求解采用隐式方案, 其他的波动则采用显式方案。第三种模式框架方案是采用半隐式半拉格朗日方案来求解动力方程组。这种方案的优点是能采用比前两种模式框架方案更大的时间步长。

WRF 的动力框架具有一定的优越性, 在前处理和所选物理过程相同的情况下, WRF 对大部分中尺度天气系统的高度场、风场、散度场、水汽通量场以及垂直速度场等物理量的模拟效果要好于 MM5。

WRF 模式系统具有可移植、易维护、可扩充、高效率、方便等诸多特性, 使新的科研成果运用于业务预报模式更为便捷。WRF 有着广泛的应用前景, 包括在天气预报、大气化学、区域气候、纯粹的模拟研究等方面的应用, 它将有助于开展针对不同类型、不同地域天气过程的高分辨率数值模拟。

(28) AUTODOCK

AUTODOCK (<http://autodock.scripps.edu/>) 是由 David S. Goodsell 博士在 Olson 实验室撰写完成。之后的版本又加入了新的科学理解策略, 计算过程更稳健和快速。AUTODOCK 采用模拟退火和遗传算法来寻找受体和配体最佳的结合位置, 用半经验的自由能计算方法来评价受体和配体之间的匹配情况。

在 AUTODOCK 中, 配体和受体之间结合能力采用能量匹配来评价。最初版本中, 能量匹配得分采用简单的基于 AMBER 力场的非键相互作用能。非键相互作用来自于三部分的贡献: 范德华相互作用、氢键相互作用以及静电相互作用。在 3.0 版中, AUTODOCK 提供了半经验的自由能计算方法来评价配体和受体之间的能量匹配。

在 3.0 版本中, AUTODOCK 采用了一种改良的遗传算法, 即拉马克遗传算法 (LGA)。测试结果表明, LGA 比传统的遗传算法比模拟退火具有更高的效率。LGA 方法把遗传算法和局部搜索 (local search) 结合在一起, 遗传算法用于全局搜索, 而局部搜索用于能量优化。在 AUTODOCK 中, 局部搜索方法是自适应的, 它可以根据当前的能量调节步长大小。LGA 算法引入了拉马克的遗传理论, LGA 最大的特点就是通过进化映射 (developmental mapping) 把基因型转化为表现型而实现局部搜索和遗传算法的结合。

(29) VMD

VMD (<http://www.ks.uiuc.edu/Research/vmd/>) 是分子可视化程序, 是一种观察和分析分子动力学模拟 (molecular dynamics simulations) 的程序, 由伊利诺斯大学研究人员研制并免费向广大用户开放。

VMD 用 OpenGL 提供高质量的 3D 分子图形, 用于显示、动画演示和分析大的生物分子体系; 能够控制大型的包括一百多万个原子的三维体系。此外, 多重芯片 (Multiple Alignment tool) 也是 VMD 一个重要组成部分。

VMD 主要有如下特性:

- 多种分子显示方式和着色方法;
- 立体显示功能;
- 包含了 Babel 程序, 用于格式转化;
- Tcl 语言编写的图形和字符用户界面, 用户可以自己编写自程序用于分子的分析;
- 用户可以用 C++ 编写模块;
- 整合了 NAMD 程序, 一起快速稳定地进行分子动力学并行计算, 并显示计算结果;
- 录制轨迹。

(30) TINKER

TINKER (<http://dasher.wustl.edu/tinker/>) 是一些程序的集合, 用于进行分子力学和分子动力学计算, 以及一些用于生物聚合物计算的特殊功能。

TINKER 具有如下功能及特色:

- 可以支持的力场: AMBER-94/96、CHARMM27、MM2 (1991)、MM3 (2000)、OPLS-AA 和 OPLS-UA;
- 包含多种算法: 新型距离 - 结构度量方法; Elber 的反应路径方法; 用于全局优化的几种势能曲面平滑和寻找 (PSS) 方法; 高效的势能曲面扫描程序; 高效简化的 Newton (TNCG) 定位优化程序; 简单的自由能微扰工具; 速率 Verlet 随机动力学; 改善的球面能量截断方法; 最新的长程静电反应场处理方法; 模拟退火; 分子动力学的 Andersen 随机碰撞恒温等;
- 经过改善的晶体最小化算法;
- 可以结合 RasMol、ChemDraw、Chem3D、gOpenMol、MOLDEN、GAMESS-US 和 ReView 等一起使用;
- PDB 文件可以转化为 TINKER 文件, 使得分子结构的输入有丰富的来源;
- TINKER 可以串行和并行方式运行。

(31) OpenMX

OpenMX (<http://www.openmx-square.org/index.html>) 是材料模拟程序包, 用于实现基于密度泛函理论的大标度从头计算。在密度泛函计算中有三个部分是相当耗时的: 求解哈密顿矩阵元素, 求解泊松方程, 以及对角化广义久期方程。而在 OpenMX 中, 根据计算量和内存, 几乎可以用 $O(N)$ 标度完成每一步。OpenMX 是涵盖生物材料和复合材料的纳米尺度材料科学中具有强大效用的工具。

OpenMX 具有如下功能:

- 用 Fourier 变换求解双中心积分;
- 用固定实空间网格求解三中心积分;
- 用 FFT (快速傅立叶变换) 求解泊松方程;
- 交换 - 相关项使用局域密度近似 (LDA, LSDA) 和广义梯度近似 (GGA);
- $O(N)$ 标度“分而治之”方法和广义的分而治之方法, 求解本征值问题;
- 计算使用赝原子轨道和赝势 (ADPACK 原子密度泛函程序计算产生)。

此外对于大标度的计算，OpenMX 用 MPI 方式进行并行计算，每个节点都是动态分配内存，实现了在分布式内存系统上的资源均衡。

(32) VENUS

VENUS (<http://homepage.mac.com/fujioizumi/visualization/VENUS.html>) 程序主要用于进行晶体结构和电子 / 核密度 (如来自 WIEN2k, GAMESS, VASP (或其图形界面 MedeA), ABINIT, Gaussian 和 SCAT 的电子结构计算) 的三维显示和操作。VENUS 还能显示 GAMESS, ABINIT, Gaussian 和 SCAT 计算的静电势和波函。

此外，VENUS 还包含两个用最大熵方法 (MEM) 的程序，用于三维显示电子密度；核密度和 Patterson 函数。

VENUS 由四个独立的程序组成：VEND 显示电子密度和核密度；VICS 显示晶体结构；PRIMA 进行迭代 MEM 分析；ALBA 进行此后的 Le Bail 分析。

VICS 可以输入 / 输出多种格式的结构数据文件和三维网格数据 (电子密度，静电势等) 文件。文件格式化工具包括：Alchemy、conrd、wien2venus.py、Cut3D 和 ELEN。

支持的图片输出格式有：BMP、EPS(像素数据)、EPS(矢量图数据)、JPEG、JPEG 2000、PPM、RAW、RGB (SGI)、TGA 和 TIFF。