

# 面向基因组数据的FPGA超算

FPGA Supercomputing System for Genomic Data



权建校

江苏微锐超算科技有限公司 技术副总

2018.10.16



关于微锐超算

1

FPGA超算  
体系结构

2

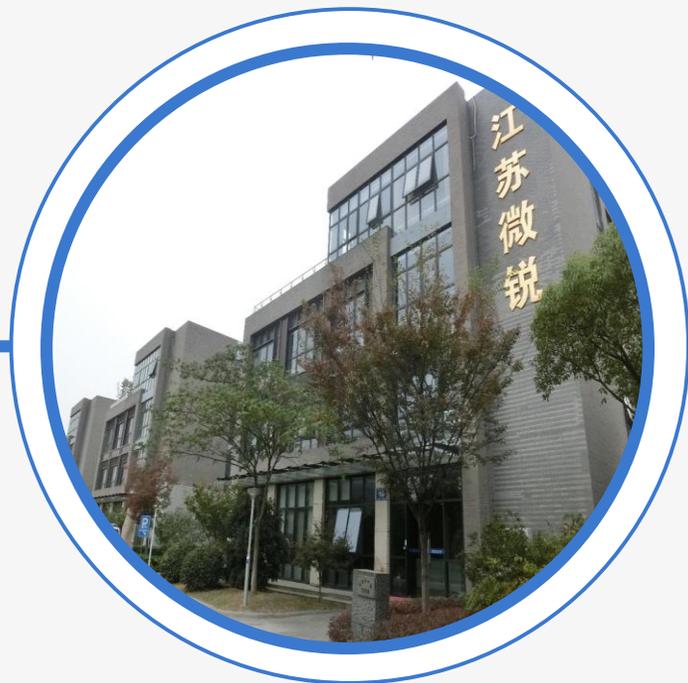
目录

FPGA超算  
应用系统

3

FPGA超算  
未来

4



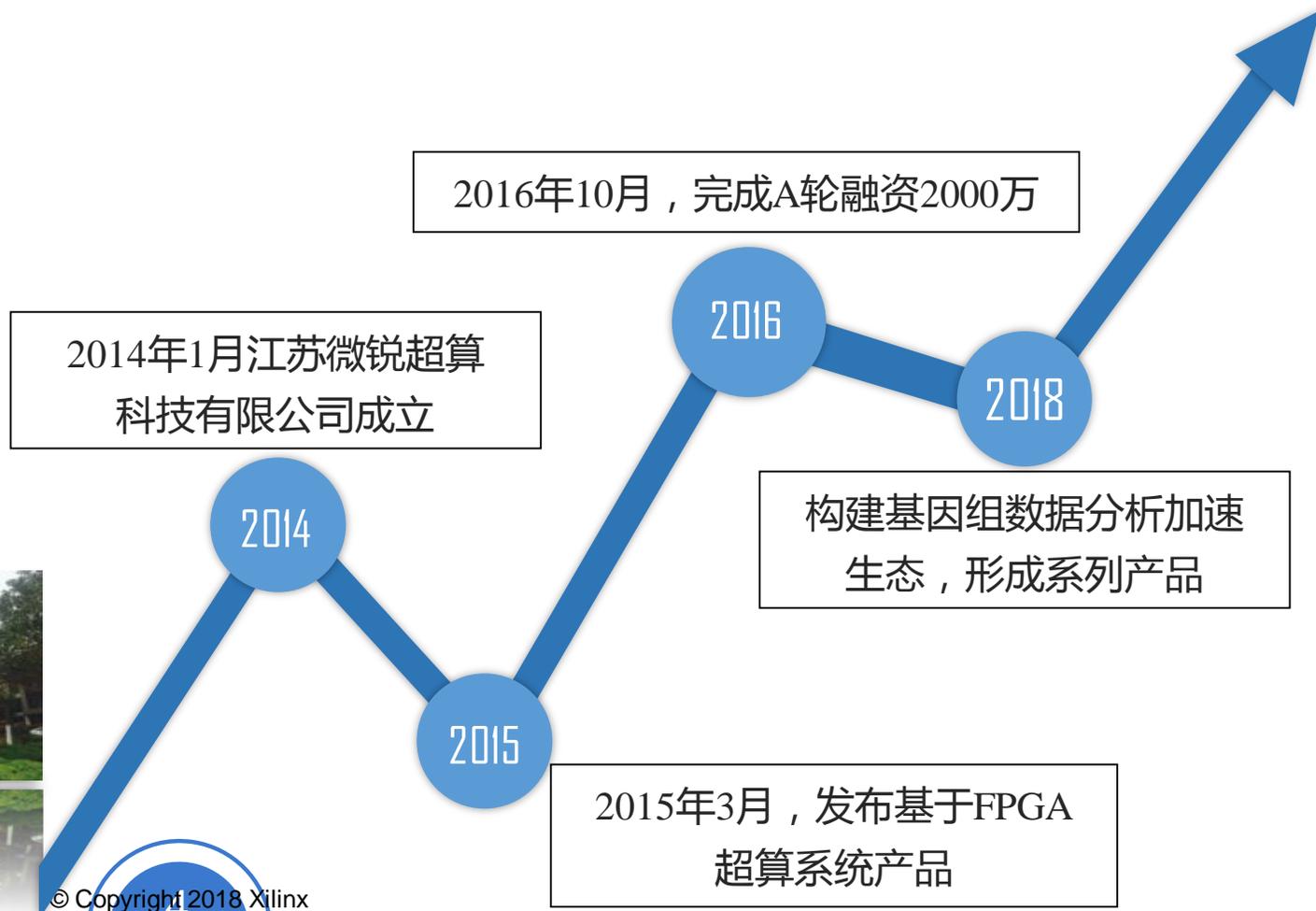
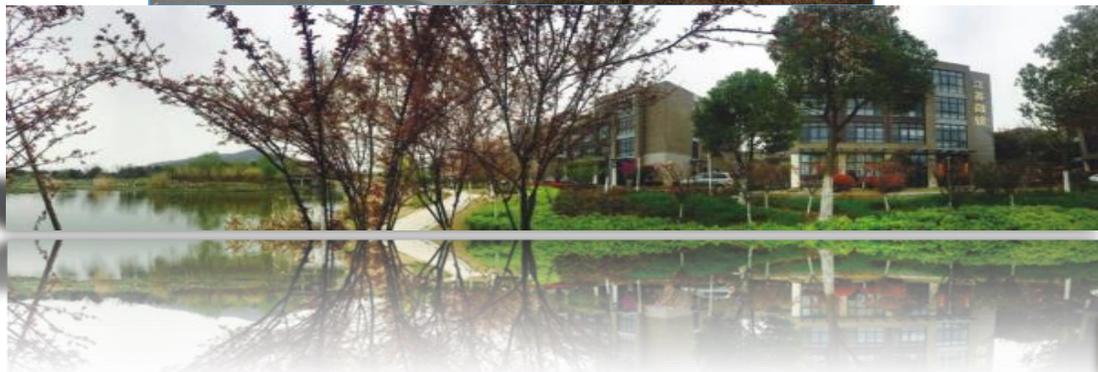
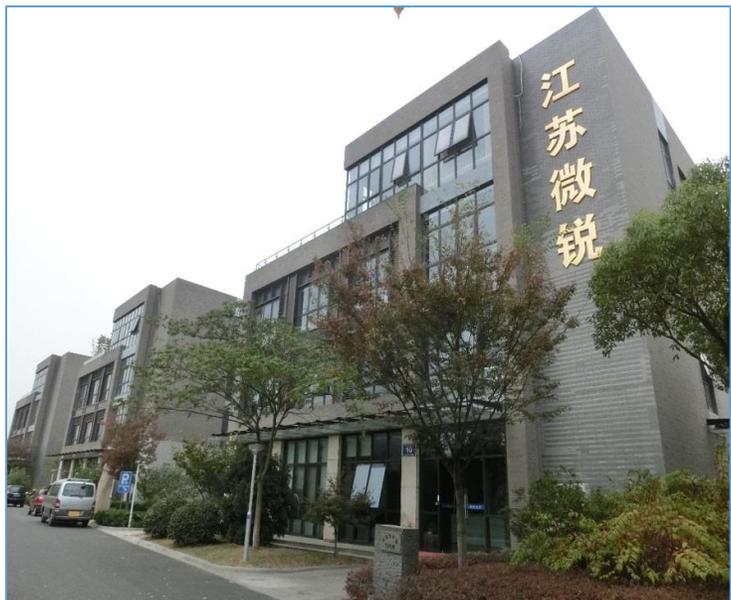
1

## 关于微锐超算



## • 江苏微锐超算科技有限公司

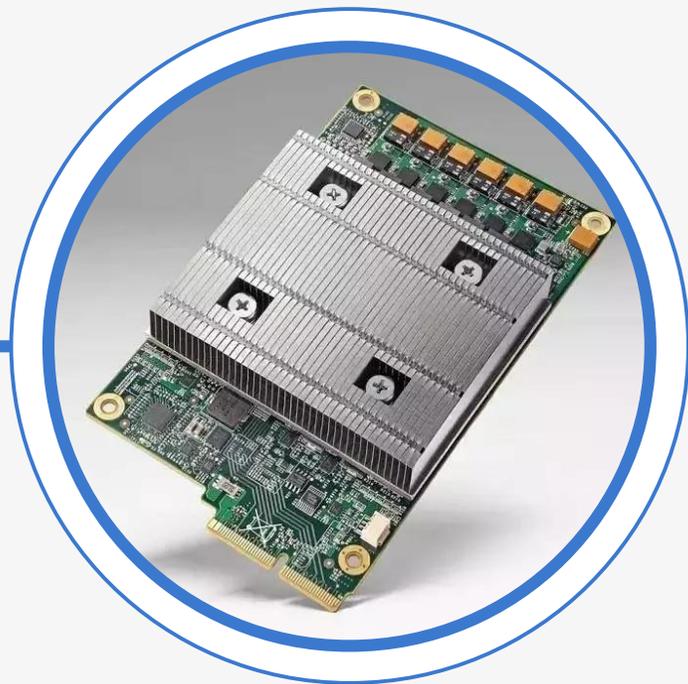
- 是一家专业从事**低功耗、高效能、高密度**FPGA超算设计、开发、销售的高科技公司。



## 02 微锐超算与FPGA

# 面向行业领域提供低功耗、高效能、高密度的超级计算系统解决方案

<p>More Flexible</p> <p>More Efficient</p>	CPU	Per/W	Today's standard, most programmable, Good for service changing rapidly.	Conventional programming
	Many-core CPU	1X	Many simple cores(10s to 100s per Chip),useful if software can be fine-grain parallel, difficult to maintain	
	GPU	3X	Good for data parallelism by merged threads(SIMD),High memory bandwidth, power hungry	Alternative programming
	<b>FPGA</b>	5~30X	Most radical fully programmable Option, Good for streaming parallelism, power efficient but program in H/W languages	<b>Alternative programming</b>
	Structured ASIC	5~100X	Lower-NRE ASICs with lower performance/efficiency. Includes domain-specific (programmable) accelerators	Can't change functionality
	Custom ASIC	20~100X	Highest efficiency. Highest NRE costs. Requires high volume, Geed for functions in very widespread use that are stable for many years	
		>100X		



2

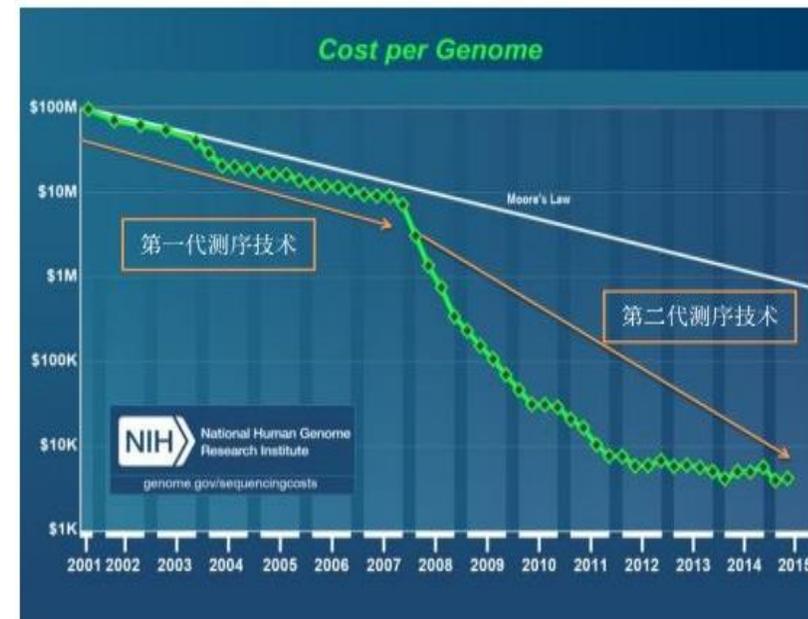
## FPGA超算硬件体系



## 01 技术背景与挑战

### 基因组大数据爆发式增长

据NCBI最新的数据统计2017年较2016年，基因组数据项目增长了**49.94%**，存储于GenBank数据库中的碱基对数量已经超过**2.6万亿**。美国冷泉港实验室的研究者认为，到2025年每年将有1ZB的基因组数据产生。



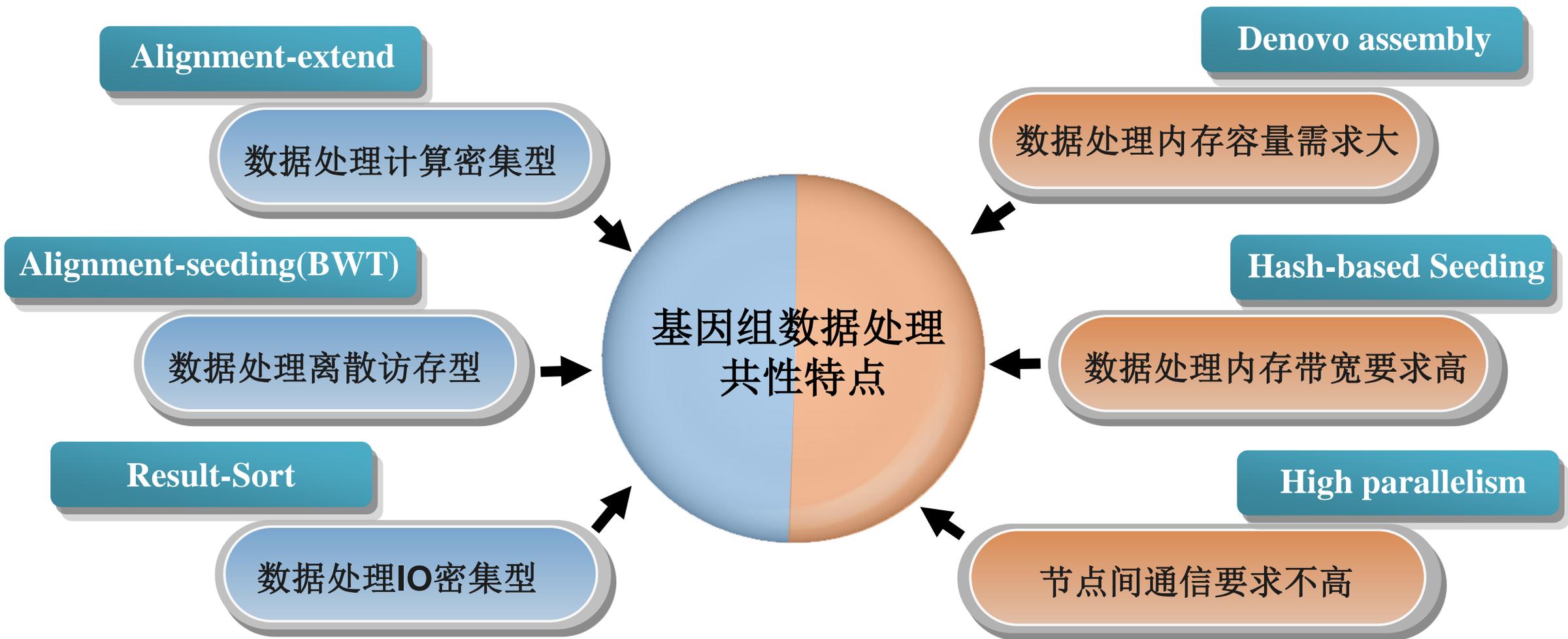
## 计算的挑战

- ◆ **数据处理效率挑战：** 数据处理算法、数据处理方法、计算机体系结构挑战等；
- ◆ **数据处理可扩展性挑战：** 数据分析过程的可并行挑战，以充分利用超算系统扩展加速计算过程

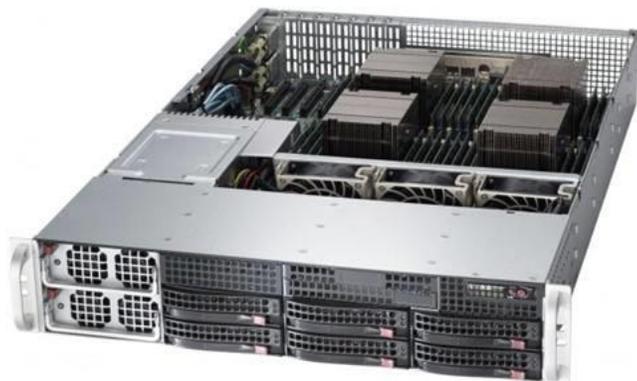
## 存储的挑战

- ◆ **数据存储挑战：** 海量数据低成本、高可靠的存储方式挑战
- ◆ **数据检索挑战：** 数据优化的检索方式及对物理存储的性能要求挑战

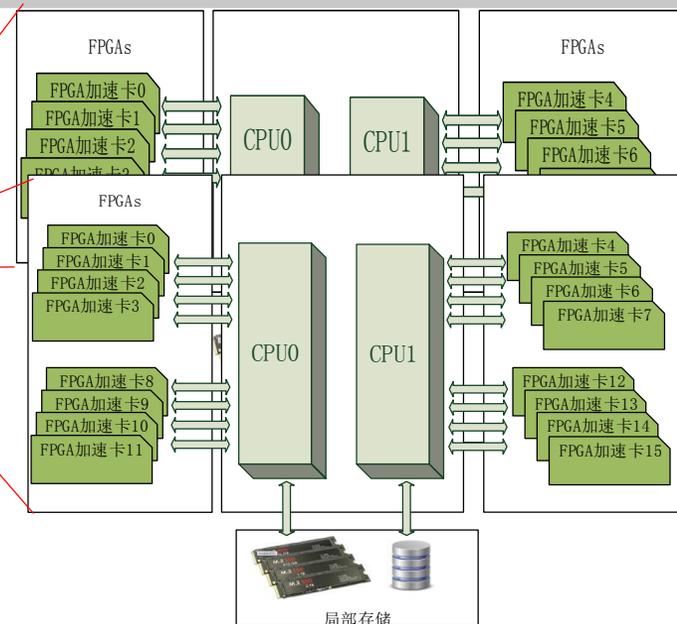
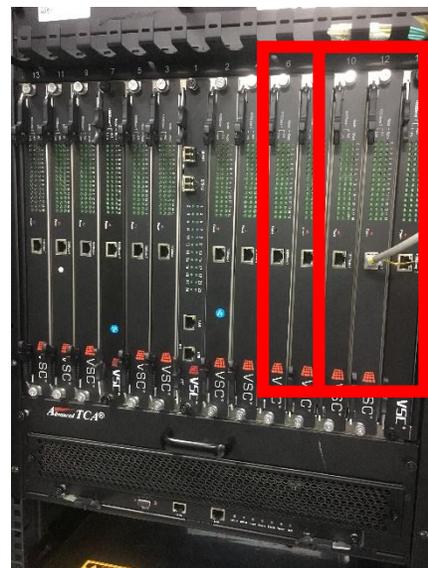
# 01 技术背景 — 基因组数据处理的共性特点



## 02 FPGA超算 —— 核心运算部件设计



标准服务器+FPGA加速卡



### 弹性可伸缩超节点特点

- 采用标准ATCA架构、计算密度高；
- 标准配置一个X86的刀片+2个加速卡刀片，集成384GB内存、1TB SSD、4TB SATA；
- 访存聚合带宽达到 691.2GB/s；
- 根据应用特点灵活配置通用算力和专用算力，典型配置为1（通用计算刀）+2（专用计算刀），并提供1+1，1+2，1+4，1+8的配置

## 系统构成

## ● 核心运算部件

弹性可伸缩超节点、胖节点

## ● 中央交换网络

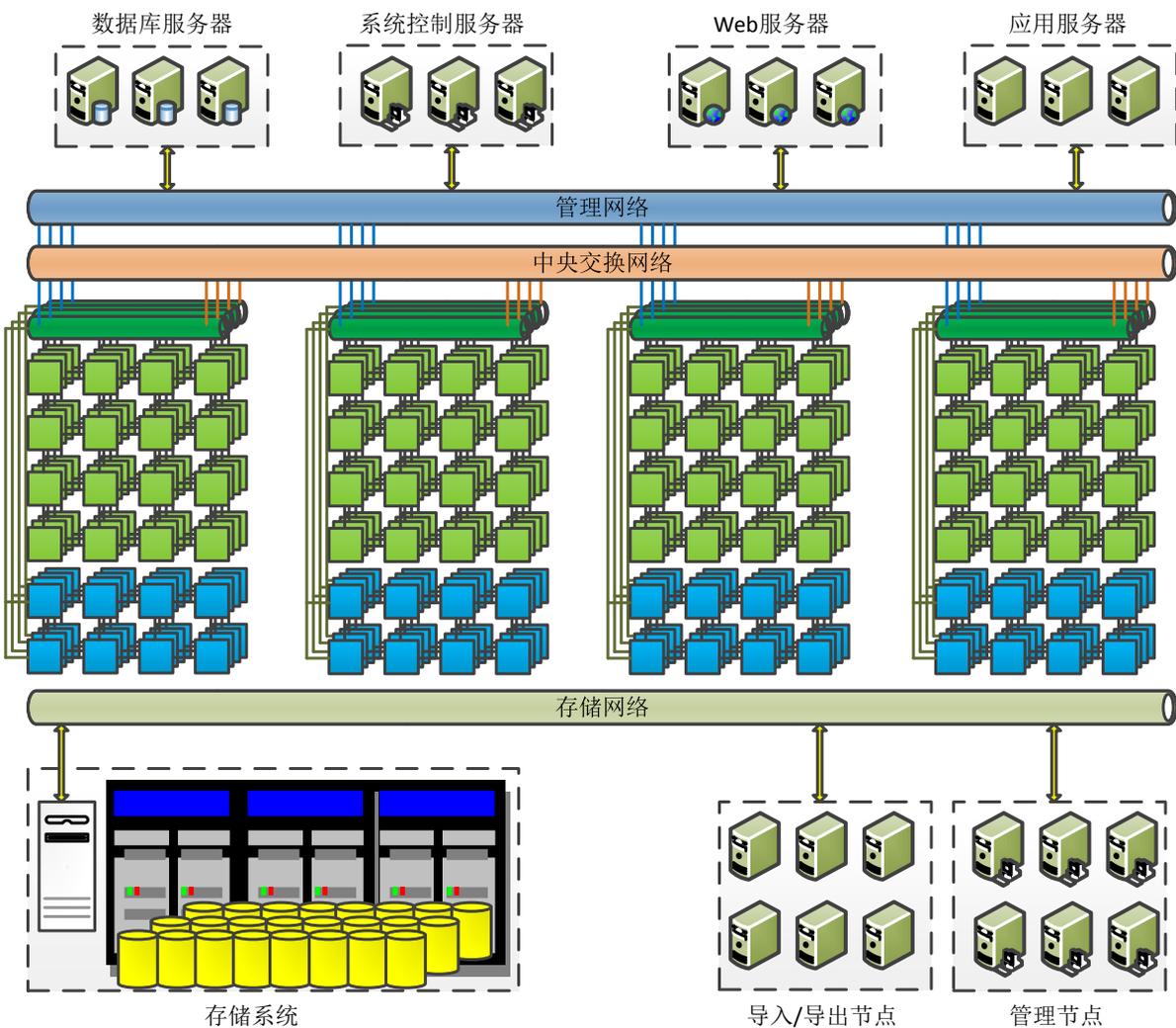
负责节点间的任务划分、数据交换

## ● 管理网络

负责节点状态的监控、任务提交等

## ● 存储系统

提供数据存储、读取、写入等



# 03 系统软件结构



分层设计，降低复杂度



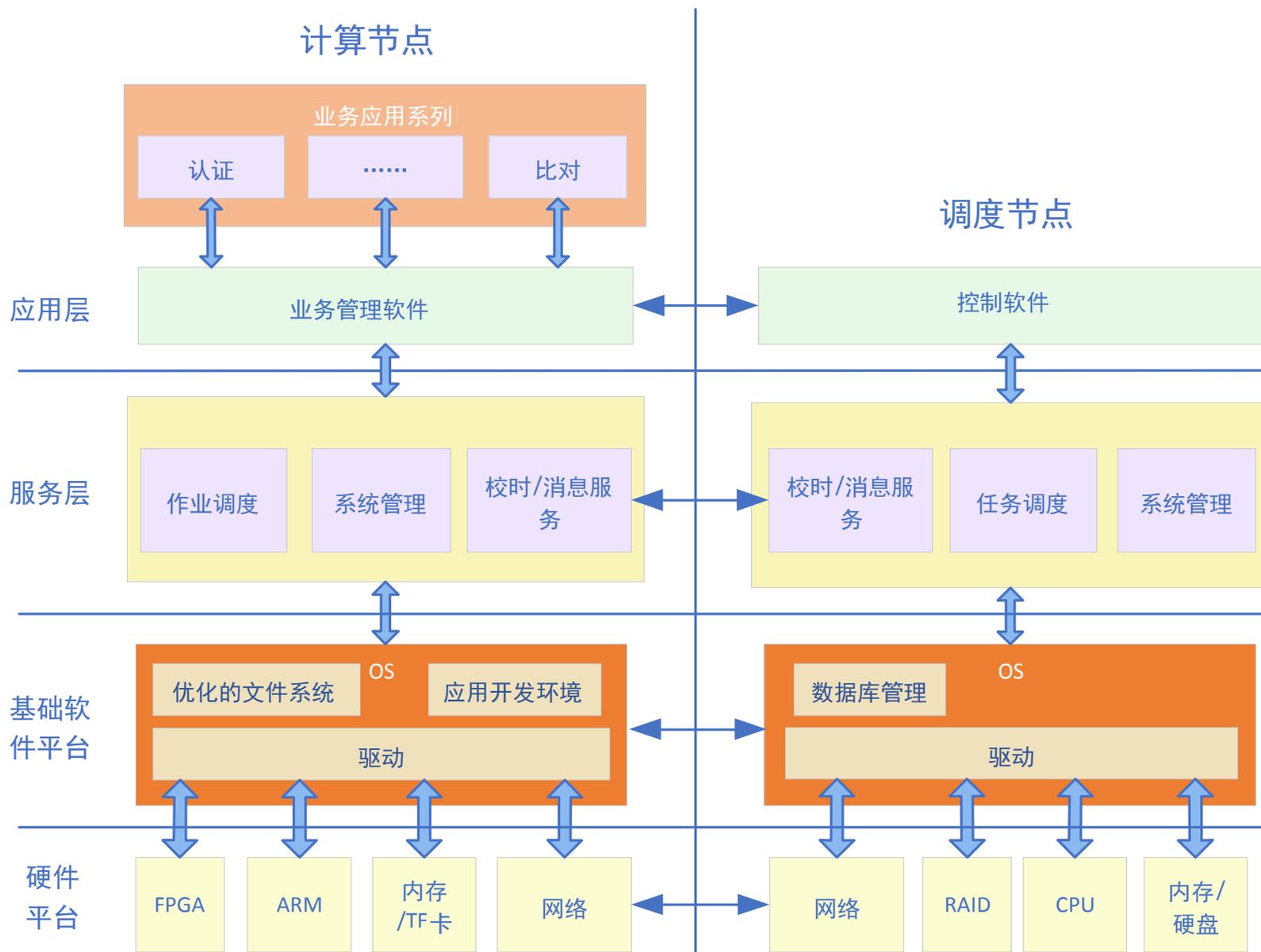
接口标准化，高可扩展



高并发调度，高加速比



容错设计，高可靠性

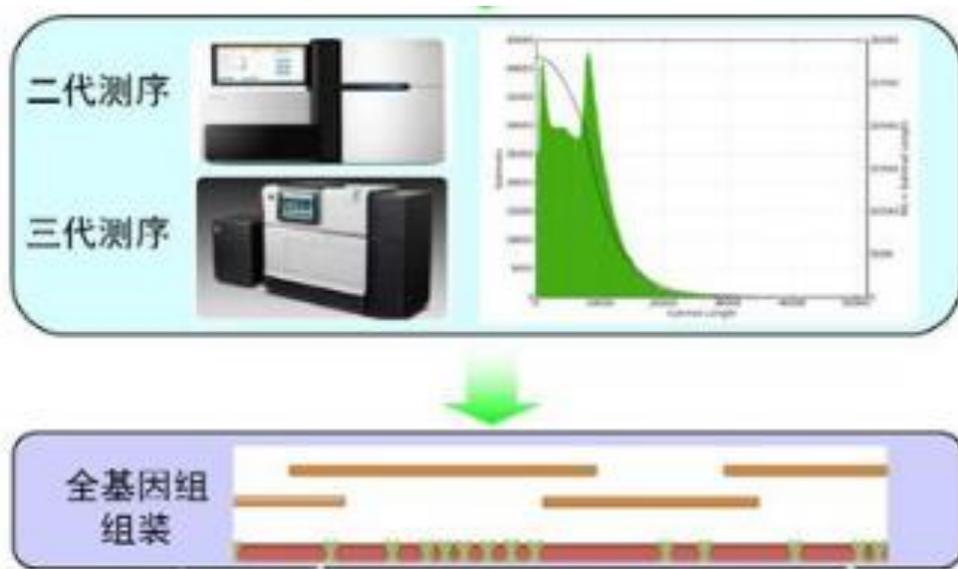




3

## FPGA超算应用系统





基因组装

三代数据组装单节点性能提升15~23倍



基因变异分析

BWA+GATK单节点性能提升12倍

## 三代单分子实时测序特点:

- 长读长平均14kbp(二代: 50-200bp)
- 无需PCR扩增, 不会引入GC偏好
- 序列错误率高 (15%)



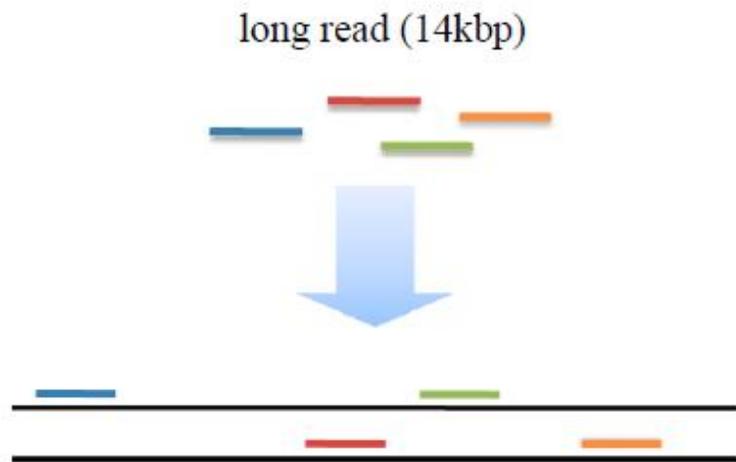
## 组装计算难题:

- 序列比对重叠区域寻找耗时

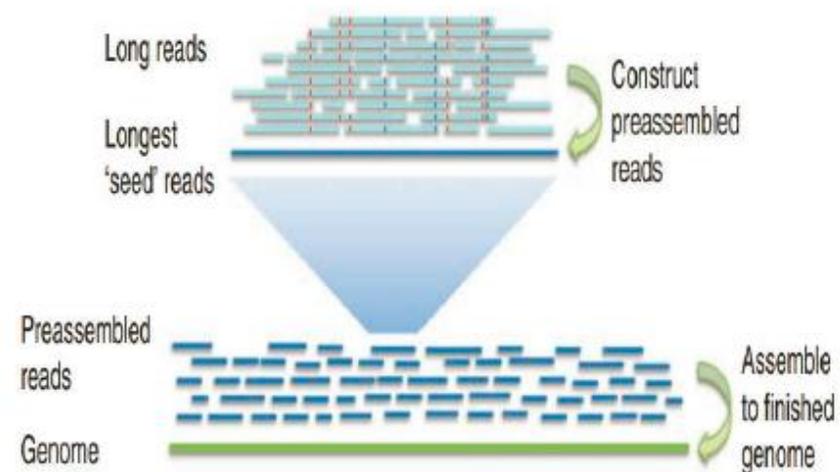
**250,000 CPU** hours to  
assemble 54x human  
--*Nat. Biotechnol* (2015)

33:6,623

## 参考基因组比对



## 两两序列比对



## 02 三代比对软件 —— Minimap2

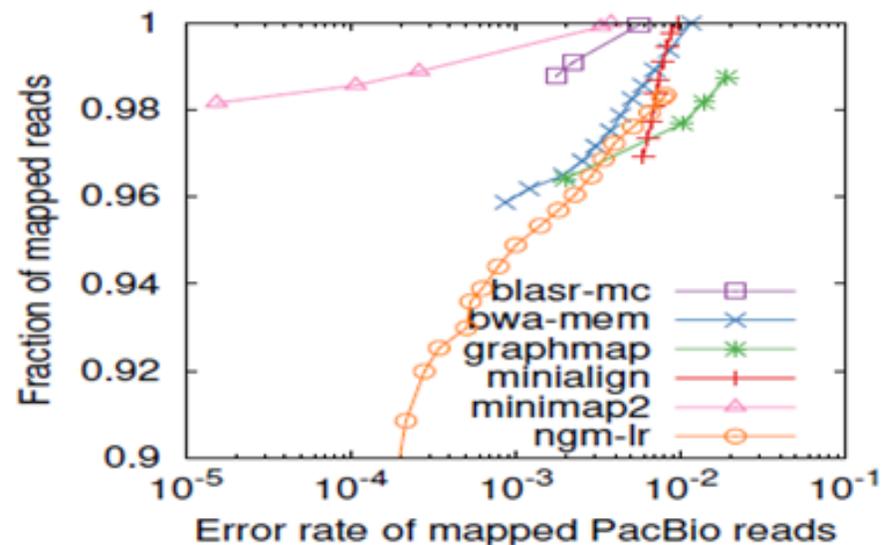
### 软件概况介绍

Minimap2是Broad Institute Li Heng开发的快速三代测序比对软件

主要应用场景：三代数据的mapping和组装Overlap查找

<https://github.com/lh3/minimap2>

软件 (40线程)	时间
minimap2	<b>128m0.008s</b>
BWA MEM	2700m59.175s
NGMLR	1576m2.239s



## 02 三代比对软件 —— Minimap2

### 特点统计与分析

#### Mapping模式

- Seeding: 基于 $(w,k)$ -minimizers建立hash table
- Chaining: 对生成的seed生成链
- Extend: 引入SSE加速SW过程

模块名称	时间	时间占比	特性
Seeding	12445s	13.56%	访存密集
Chaining	3495s	3.81%	计算&访存密集
Extending	74491s	81.17%	计算密集

测试环境: 双路E5-2680 V4处理器

#### AVA模式

- Seed: 基于 $(w,k)$ -minimizers建立hash table
- Chaining: 对生成的seed生成链

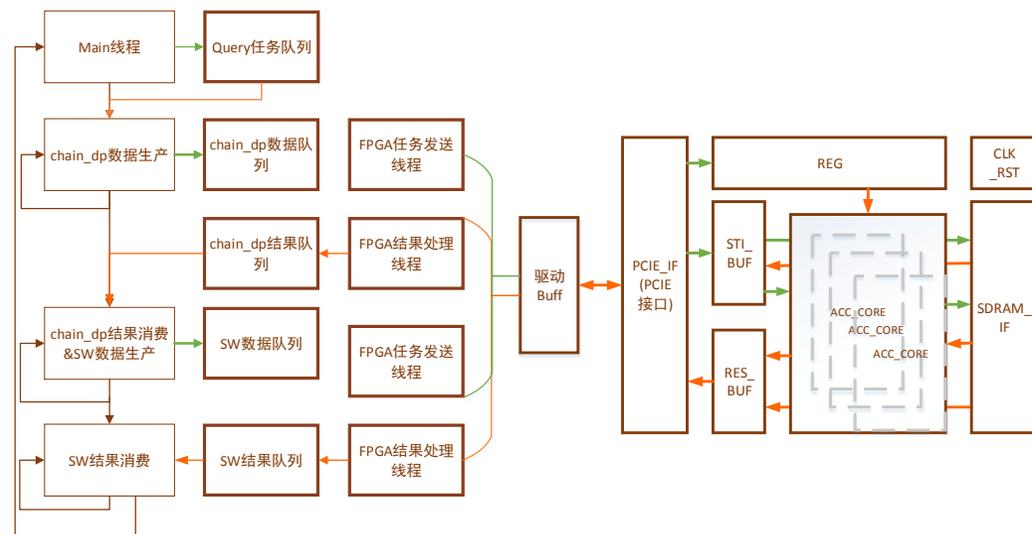
模块名称	时间	时间占比
Seeding	527248s	67.5%
Chaining	239147s	30%

测试环境: 双路E5-2680 V4处理器

## 02 三代比对软件 —— Minimap2-hp

### 优化方案及效果

- 优化核心数据结构，减少数据访问离散度；
- 优化数据访问方式，以局部计算替代访存；
- 重构软件架构形成“任务调度—核心模块输入输出—加速器运算处理”完全流水的基本架构；
- 算法核心运算采用加速器实现；



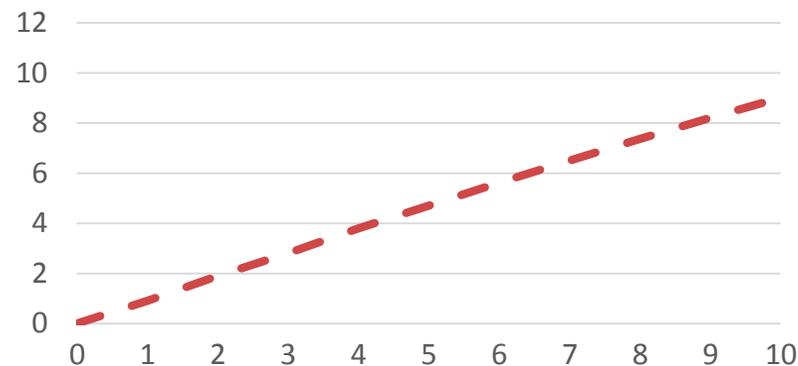
AVA模式下优化效果测试

数据集	Minimap2	Minimap2-hp	Speedup
np.new.fasta	1579s	75s	20.05
huaxia.15X.fasta	16810s	684s	24.57

Minimap2: 双路E5-2680 V4处理器 56线程

Minimap2-hp: 双路Gloden5120处理器, 8张加速卡

Minimap2-hp线性加速比



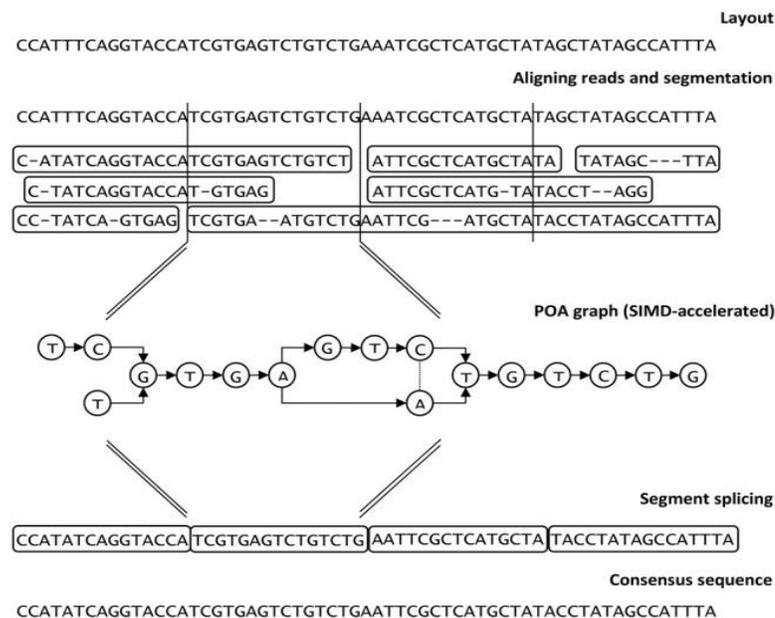
## 02 三代序列校正软件 —— Racon

### 软件概况介绍

Racon是一款对三代测序组装结果以及原始三代测序数据进行校正的软件

<https://github.com/isovic/racon>

CPU时间 (min) / 内存 (GB)



基于POA图的碱基校正

	miniasm+Racon	Canu	FALCON
<i>E.coli</i> ONT 54x	<b>25/3.32</b>	1328/4.03	829/12.29
<i>E.Coli</i> PacBio 160x	<b>86/9.69</b>	773/3.60	2908/9.93

组装一致性

	miniasm+Racon	Canu	FALCON
<i>E.coli</i> ONT 54x	<b>99.13%</b>	99.28%	98.84%
<i>E.Coli</i> PacBio 160x	<b>99.63%</b>	99.99%	99.90%

## 02 三代序列矫正软件 —— Racon

### 特点统计与分析

- EdlibAlign: 计算密集型, 实现Overlap粗比对;
- SPOA::graph: 访存密集型, 构建全局图;
- SPOA::alignment: 计算密集型, 精确匹配;
- SPOA::generate\_consensus: 访存密集型, 实现对read进行校正;

软件模块运行耗时占比

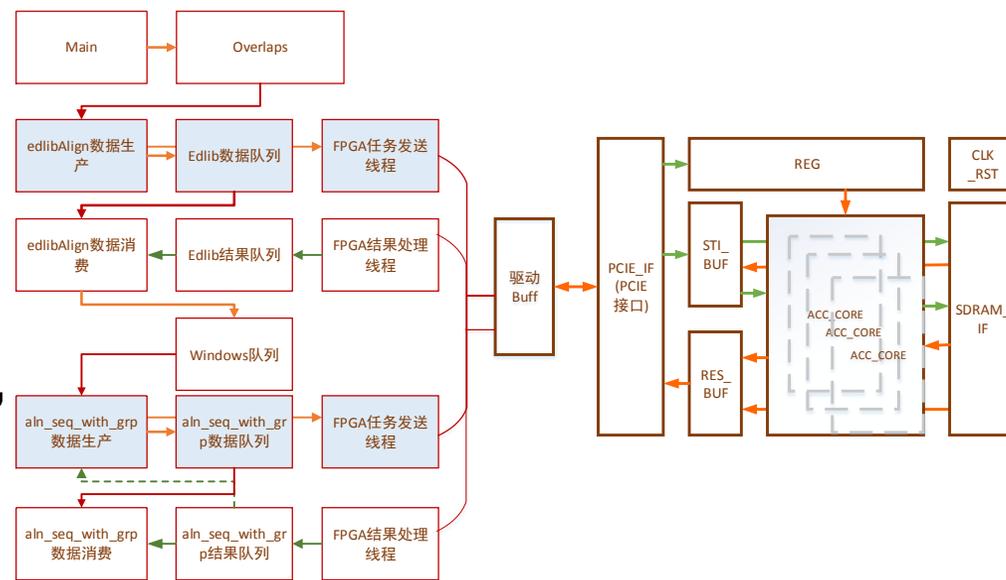
模块	测试时长	时间占比
EdlibAlign	107764s	33.6%
SPOA::graph	60677s	20.12%
SPOA::alignment	124580s	41.31%
SPOA::generate_consensus	6409s	2.1%

测试环境: 双路E5-2680 V4处理器

## 02 三代序列矫正软件 —— Racon-hp

### 优化方案及效果

- 优化软件基本结构，解决多线程计算能不能充分发挥的问题
  - EdlibAlign的操作与generate\_consensus串行处理，导致多线程加速比6；
- 重构软件架构形成“任务调度—核心模块输入输出—加速器运算处理”完全流水的基本架构；
- 算法核心运算采用加速器实现；

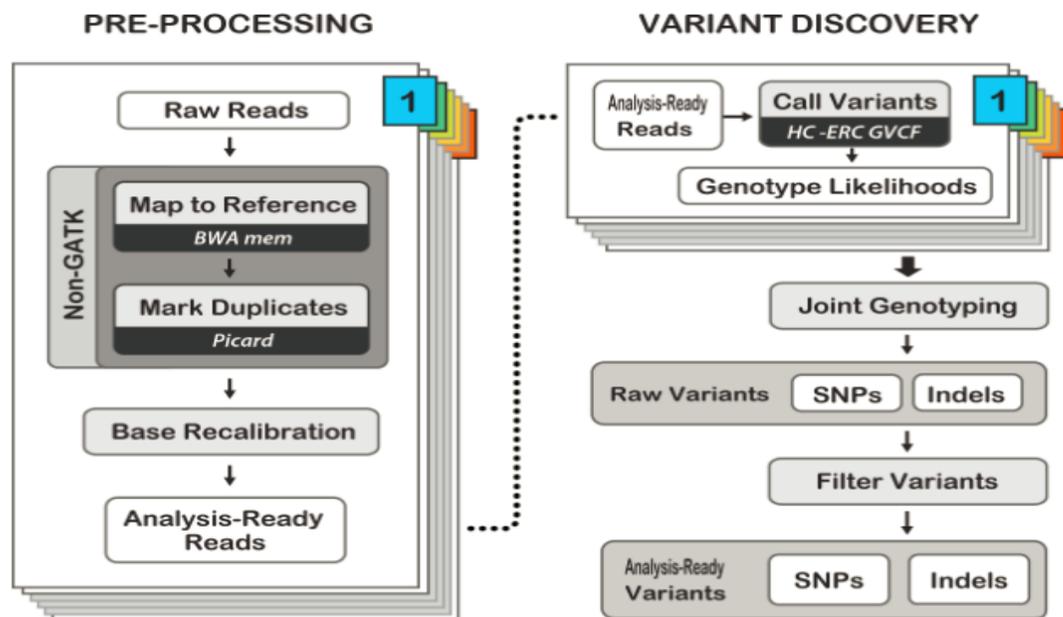


测试数据	Racon	Racon-hp	speedup
Y.pestis	34m1s	8m	4.25
B.anthraxis	211m31s	42m	5.02
C.elegans	314m42s	69m	4.5

测试环境说明

通用服务器：双路E5-2680 V4处理器 56线程

FPGA超算单节点：包含双路Gloden5120处理器，1张加速卡



Best practices for germline SNPs and InDels in WGS

Tools	1 Thread	36 Threads
BWA MEM	92h03m19s	3h50m58s
Picard SortSam	7h55m51s	6h36m35s
Picard MarkDuplicates	6h34m47s	5h45m15s
GATK BaseRecalibrator	19h49m07s	1h56m07s
GATK PrintReads	23h54m38s	7h28m08s
GATK HaplotypeCaller	63h39m09s	6h18m39s
<b>Total Time</b>	<b>213h56m51s</b>	<b>31h55m42s</b>

CPU: E5 @2.3GHz; sample: 30X WGS

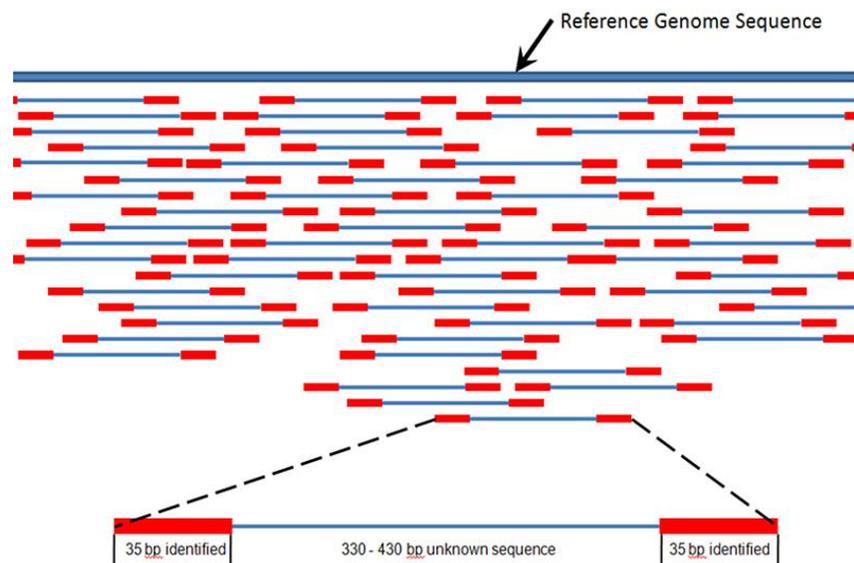
通用处理器处理基因数据时间分布

## 03 二代序列比对软件 —— BWA

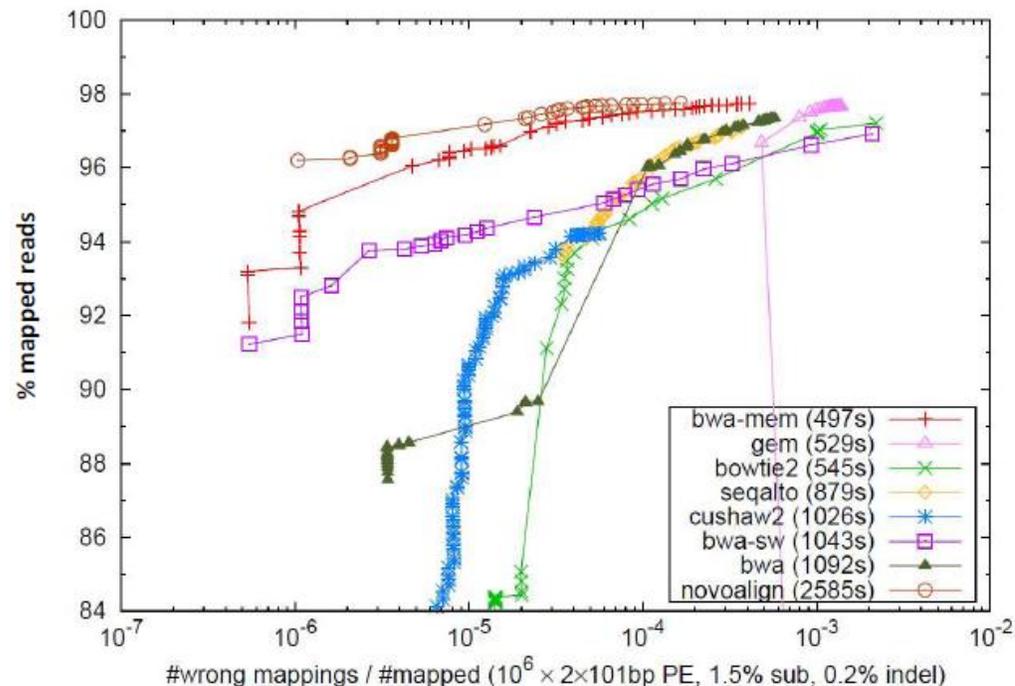
### 软件概况介绍

BWA是Broad Institute Li Heng开发的二代测序比对软件

<https://github.com/lh3/bwa>



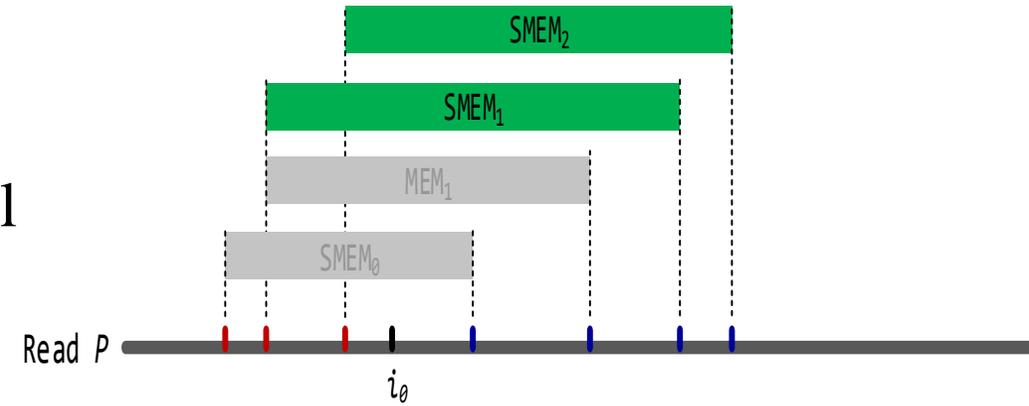
- seed: 采用FMD-Index索引查找
- extend: SW得分矩阵计算



From: Heng Li, <https://arxiv.org/pdf/1303.3997v2.pdf>

## 特点统计与分析

- Seeding: 寻找最大的SMEM (Super Maximal Exact Match)
  - 强离散访存密集型, 访存粒度小离散度大

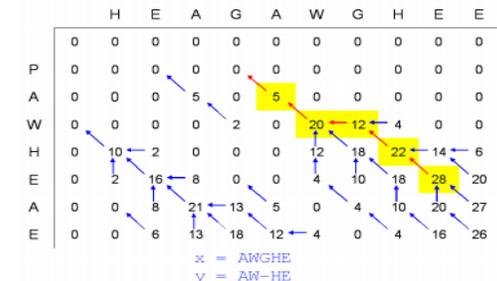


If we want to find **BEST** local alignment...

- Find  $Score_{opt}$  and then traceback

$$Score_{opt} = \max_{i,j=1}^N M[i][j]$$

- Extending: 局部比对算法对SMEM筛查
  - 计算密集型, 启发式算法



## 二代序列比对软件 —— BWA-hp

### 优化方案及效果

- 优化索引数据文件格式，压缩访存粒度；
- 优化放置索引结构，提高数据局部访问概率；
- 重构软件架构，形成异步流水计算模式；
- 将bwt\_smem1a及ksw\_extend等核心模块，采用加速器方式实现。

数据集	数据量	speedup	符合度
ERR2303730	30GB	13.22	99.88%
SRR6914852	40GB	12.11	99.67%
SRR6882893	26GB	10.53	99.94%
ERR1775135	1.1GB	12.23	99.77%

测试环境说明

BWA：双路E5-2680 V4处理器 56线程

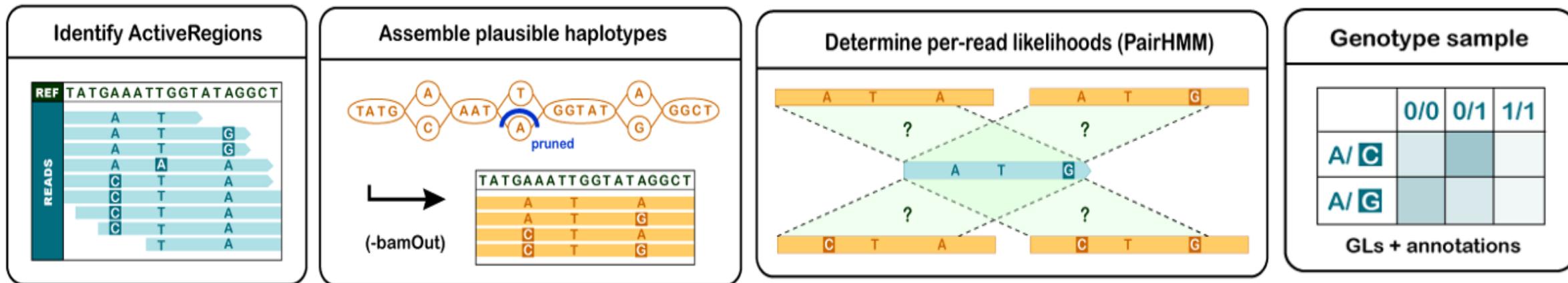
BWA-hp：双路Gloden5120处理器，8张加速卡

# 03 二代变异位点分析软件 —— GATK-HaplotypeCaller

## 软件概况介绍

GATK是Broad Insitute团队开发二代测序变异检测软件

<https://software.broadinstitute.org/gatk/>



## 特点统计与分析

- Identify ActiveRegions: IO和访存密集型
- Assemble plausible haplotypes: 访存密集型
- Determine per-read likelihoods(PairHMM): 计算密集型
- Genotype sample: 计算和访存密集型

测试输入	功能步骤	测试时间	时间占比
100G.bam	Identify ActiveRegions	11659s	4.9%
	Assemble plausible haplotypes	61670s	26.23%
	Determine per-read likelihoods(PairHMM)	145336s	61.81%
	Genotype sample	16454s	7%
	总耗时	250m52s	--

测试环境: 双路E5-2680 V4处理器

## 二代变异位点分析软件 —— GATK-hp优化方案及效果

### 优化方案及效果

- 优化数据读取与数据调度方式；
- 重构软件架构，形成异步流水计算模式；
- 将核心运算PairHMM及部分Smith-Waterman采用加速卡方式进行处理。

数据集	GATK-hp	speedup	变异位点数	变异位点 ADD	变异位点 DEL	符合度
SRR3670051	4m3s	20.35	46569	0	7	99.98%
SRR3670054	4m8s	19.83	44210	1	9	99.97%
SRR3670088	4m59s	18.82	47141	0	2	99.99%
SRR3670058	4m9s	18.99	44007	3	8	99.97%

#### 测试环境说明

GATK：双路E5-2680 V4处理器 56线程

GATK-hp：包含双路Gloden5120处理器，1张加速芯片



4

## FPGA超算的未来



## 01 FPGA将助推生命科学技术的发展



### 高端定制计算

- ◆ 针对应用需求，定制强大算力的专用超算，助力生命科学技术发展



### 敏捷边缘计算

- ◆ 通过集成高效计算部件，增强移动端的数据处理能力

# Thanks for your attention

江苏微锐超算科技有限公司

地址：江苏省无锡市滨湖区棟泽路30号  
山水城科教软件园C区10号楼

TEL : 86-510-85321179

FAX : 86-510-85321100

Official Web: [www.smart-vsc.com](http://www.smart-vsc.com)

