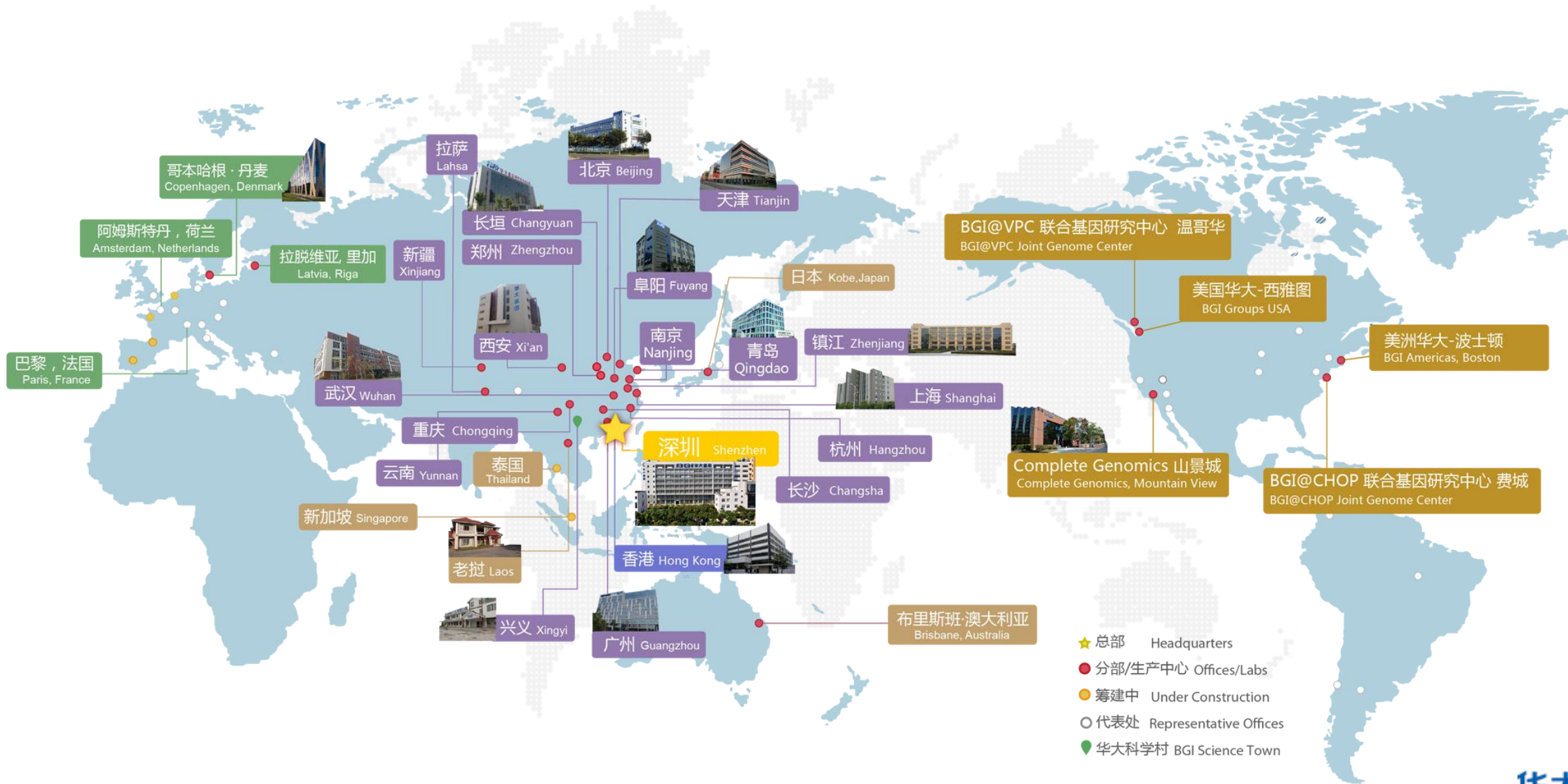




# 基因行业高性能计算系统全球化建设与运维实践

华大生命科学研究院  
刘曜玮

# 华大基因全球布局概览

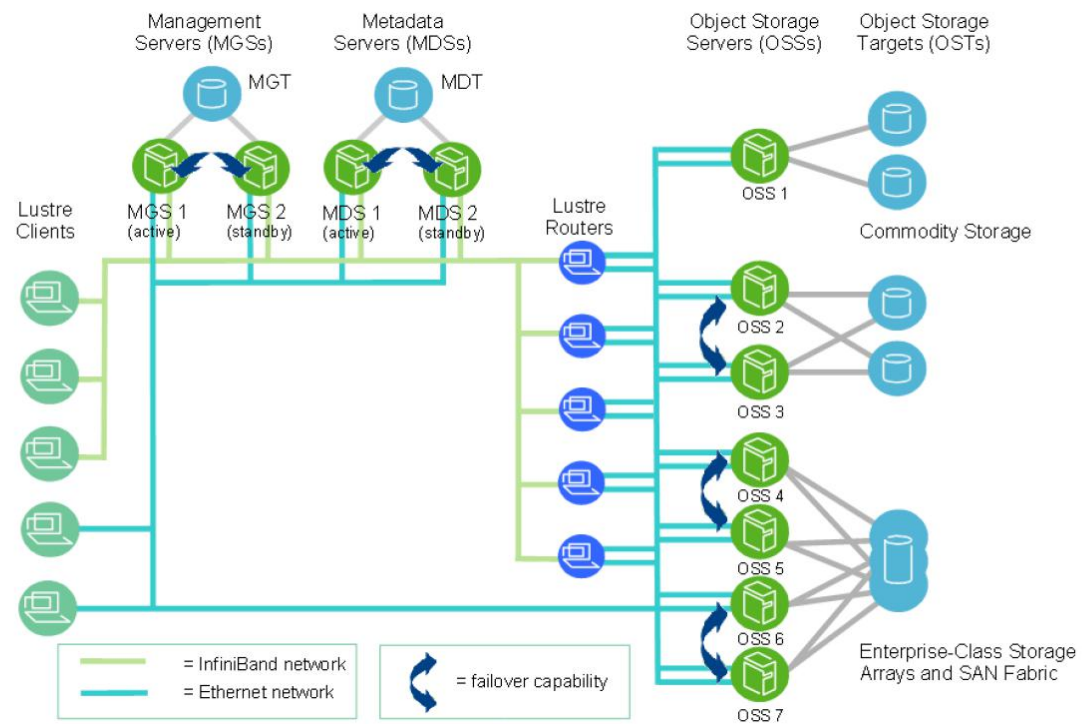


# 华大基因高性能计算系统部署区域

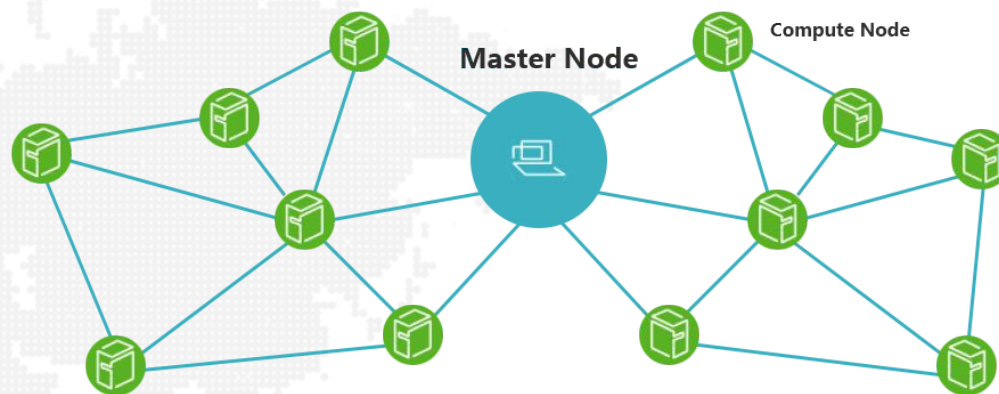
国内	深圳盐田，深圳国家基因库 武汉 香港 天津 青岛基因库 北京，河南长垣，重庆，昆明
海外	美国费城，美国萨克拉门托 丹麦哥本哈根 澳大利亚布里斯班

# 华大基因高性能计算系统介绍

## • 分布式存储



## • 网格计算系统



## • 生物信息应用与数据库



# 华大基因高性能计算系统快速建设与低成本运营

产品化，标准化，规范化，模块化

## 01 标准化

OPTION

## 02 规范化

OPTION

## 04 模块化

OPTION

## 03 产品化

OPTION



高性能计算系统在多个区域快速建设与高效运维之关键在于系统的标准化、规范化、模块化与产品化。

华大基因高性能计算系统进行了多次架构设计调整，亦经过了多个版本的迭代。根据实际运营过程中的经验，在设计过程中坚持产品化思路，以标准化、规范化与模块化为目标，从而保证高度可复制性与建设高效性。

整套系统体系架构与设计时采用分模块产品化思路，整套系统分为网格计算模块、分布式存储模块、生物信息应用与数据库模块、监控模块、日志模块、配置管理模块等。

# 华大基因高性能计算系统快速建设与低成本运营

开源产品为主，商业产品为辅

## 思路：

- 1、多区域部署，使用开源产品有效降低建设成本。规模化后，成本效益更加明显。
- 2、基因测序与生物信息行业变化与发展迅猛。开源产品具备更好的灵活性、快速迭代性。故能更好的适应不同的业务建设场景，保证建设效率。
- 3、部分关键以及特性稳定的业务采用商业产品保障。

网格计算	SUN GRID ENGINE调度系统	开源产品，深度定制调度策略
分布式存储	Lustre EMC Isilon, 华为Ocean Store, Panasas, Dell FS	开源产品 + 商业产品
生物信息应用与数据库	数百种常用生物信息软件，模块与库 NCBI, CNGB, EBI等常用生物信息数据库	行业开源软件与公共数据库

# 华大基因高性能计算系统建设

业务导向性，业务一致性，业务扩展性



- **业务扩展性**

基因行业发展迅速，系统设计与架构、建设时，需保证扩展简单高效，扩展成本可控。

- **业务一致性**

系统设计、建设均需与业务需求保持一致。尤其是建设初期，需合理评估业务量（如测序数据产量，数据分析计算需求量，数据产量...），控制成本。

- **业务导向性**

系统在设计及架构，技术方案选型，均以服务实际业务（基因测序，数据分析，数据解读...）为导向。

1：了解业务

2：理解业务

# 华大基因高性能计算系统建设

## 评估与建设落地

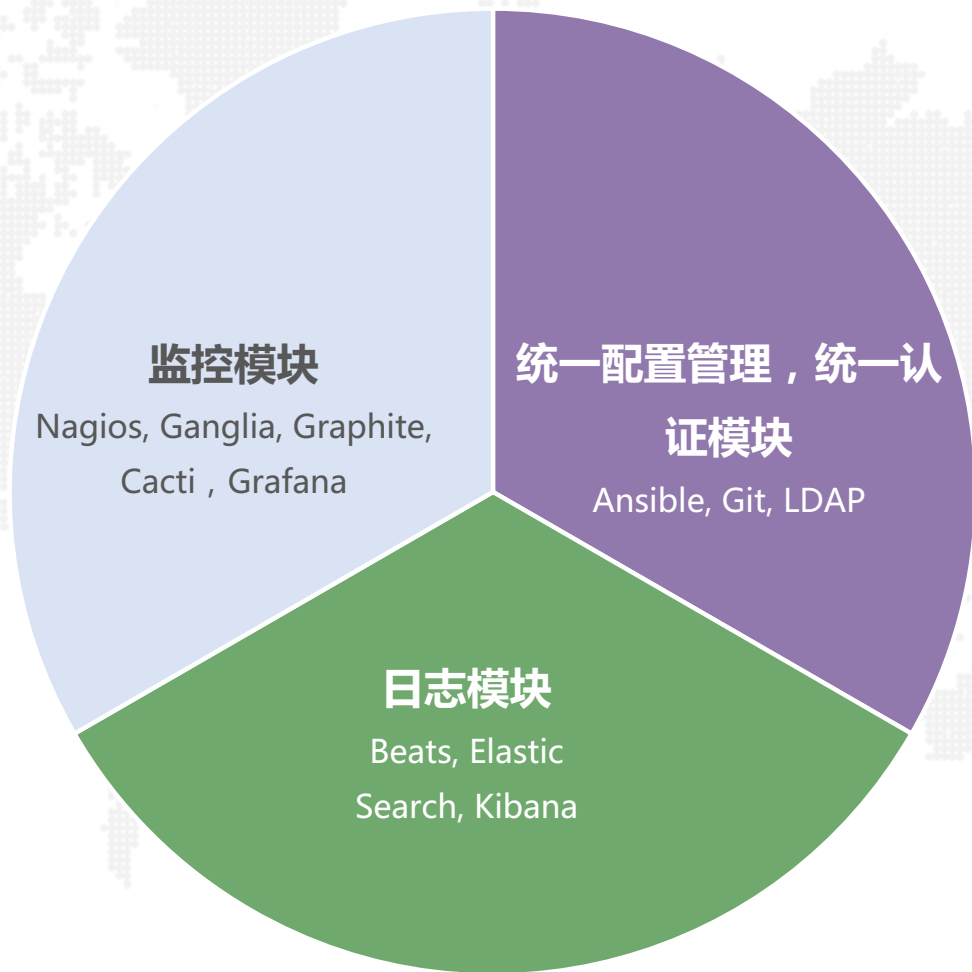


基础设施	数据中心基础设施是保证后期系统稳定运行的基石。根据业务需求评估与设计基础设施方案，合理实施基础设施建设或改造（自建+租用）。	需要关注：用电、UPS、制冷、承重、园区及电信运营商网络等
设备资源	根据业务需求、成本预算框架，参考累计的设备测试与设备选型库，选择使用的系统架构模块及设备类型。	计算、存储、网络设备资源评估与选型。除成本外，需关注：远程可维护性、密度、能耗，厂商支持与服务。
人力资源	订立建设运维与人力资源标准，根据标准评估建设与运维所需人力资源。	人力资源除自身人力资源外，也可借助外部资源，如运营商人力资源。



# 华大基因高性能计算系统全球化运维

开源技术模块化集成+定制



# 华大基因高性能计算系统全球化运维

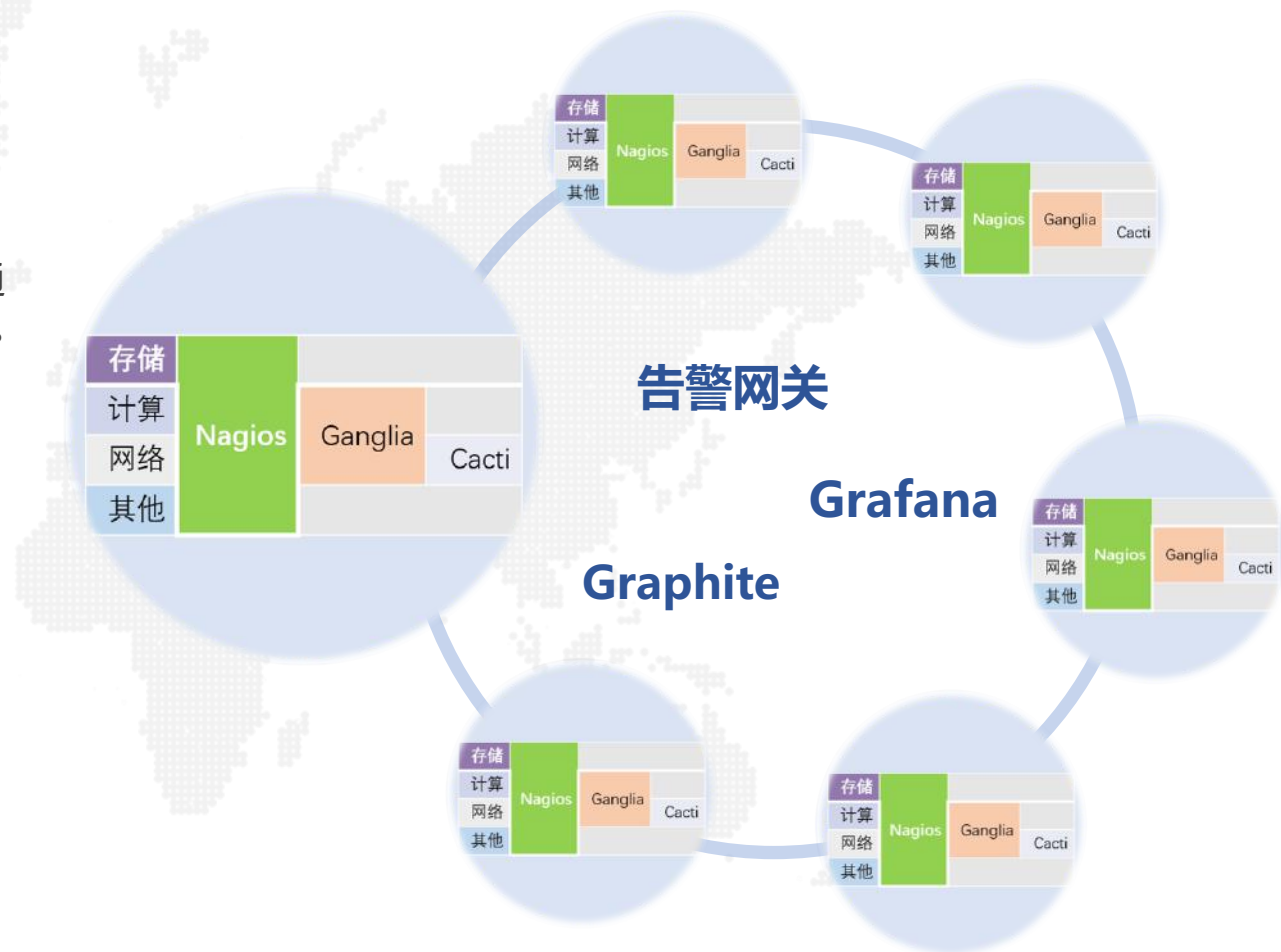
## 监控模块

- Nagios :

存储：需针对数家厂商，不同芯片，数十种存储产品（普通存储，盘柜，JBOD，商业存储等），定制监控脚本与程序。

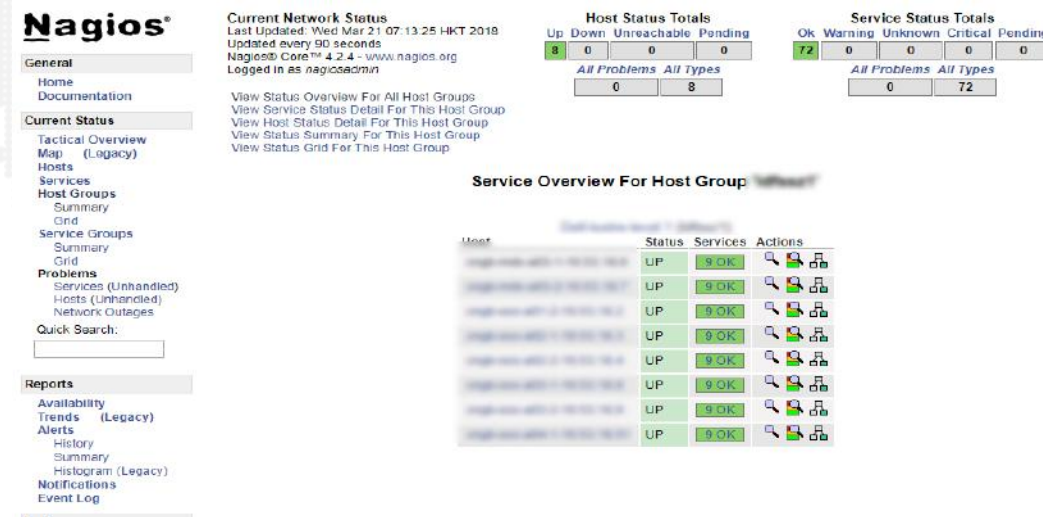
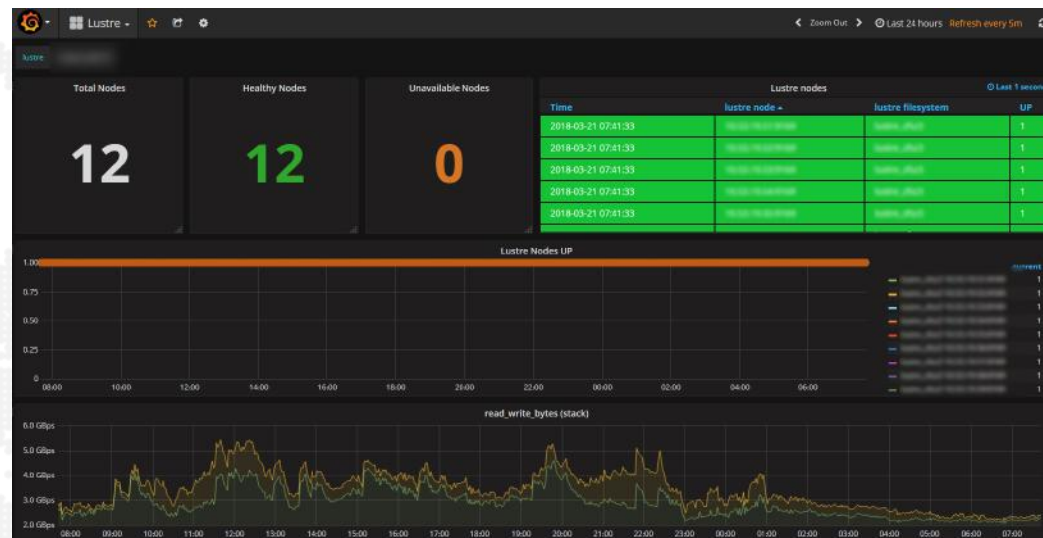
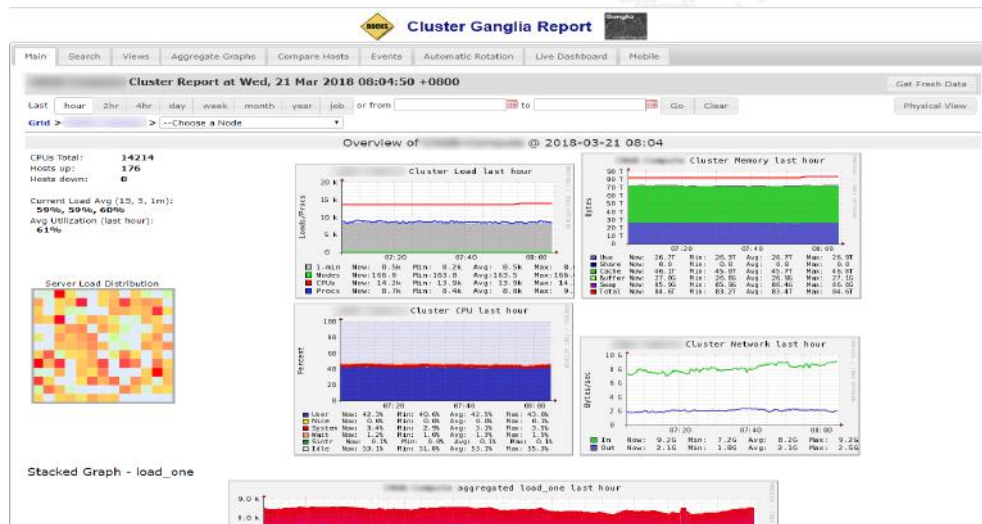
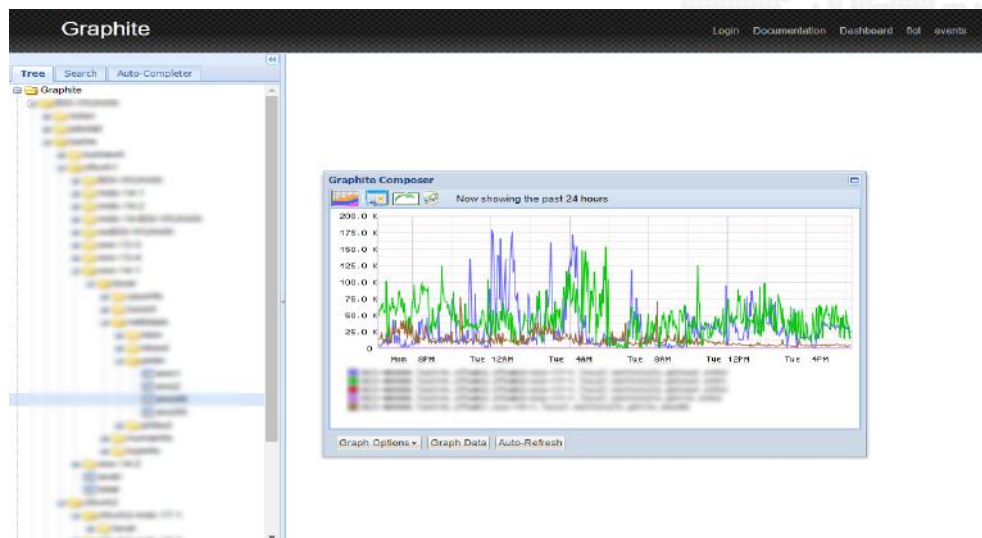
其他：主要通过IPMI, redfish接口监控温度，能耗等信息。

告警：通过告警网关发送邮件，短息与电话告警。



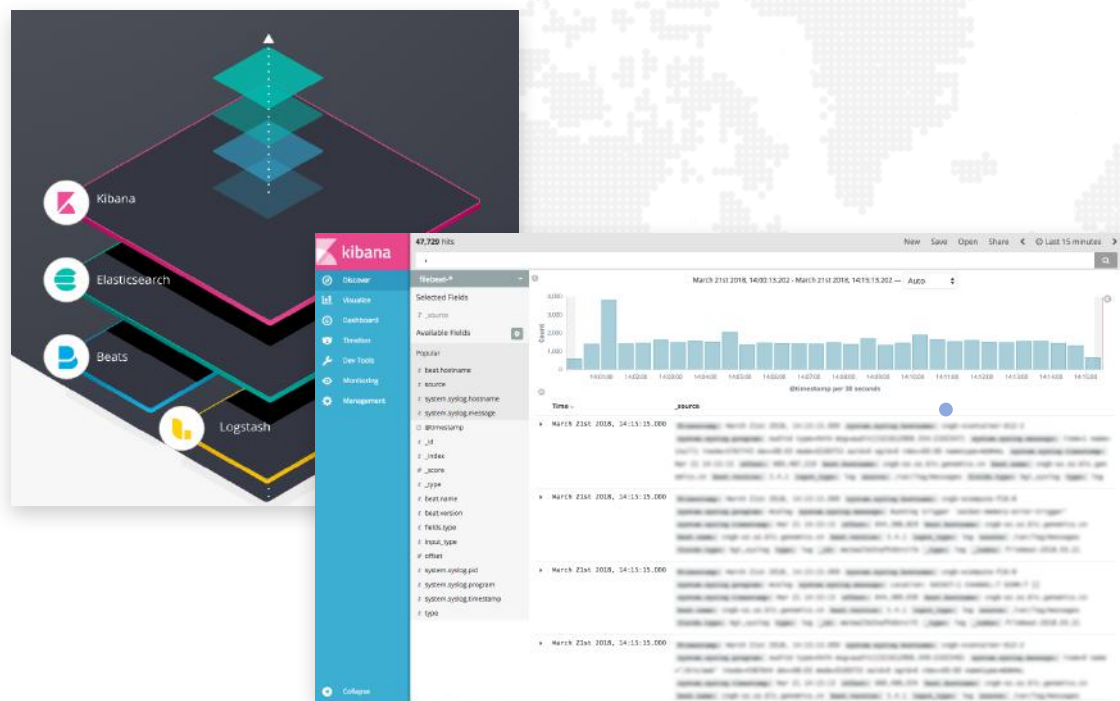
# 华大基因高性能计算系统全球化运维

## 监控模块



# 华大基因高性能计算系统全球化运维

## 日志模块



日志模块主要采用开源Elastic Stack。

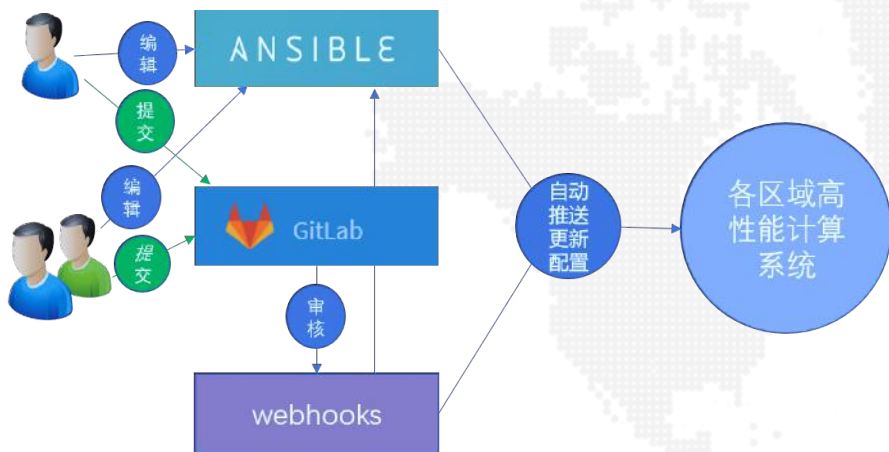
- 1、Beats --- 日志采集
- 2、Elastic Search --- 搜索，分析与存储
- 3、Kibana --- 数据可视化

### 重点关注：

- 1、调度系统日志，包括调度系统job，job运行数据调度系统日志主要用于后期的智能化运维，调度系统调度策略智能化。
- 2、分布式存储日志  
分布式存储日志主要用于后期的存储智能HSM

# 华大基因高性能计算系统全球化运维

统一配置管理，统一认证

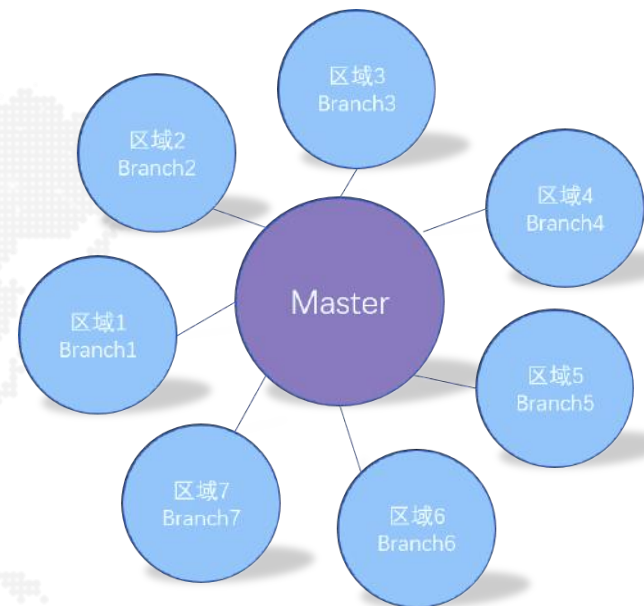


## Ansible :

DEVOPS理念保障系统快速部署与交付，并保障全区域核心配置统一管理。

## GitLab :

Git版本与协同管理理念，实现配置的多用户贡献以及版本管理。Git分支理念保证各区域系统配置灵活性。webhooks保证配置变更的审核流机制以及配置快速推送与下发。



# OpenLDAP

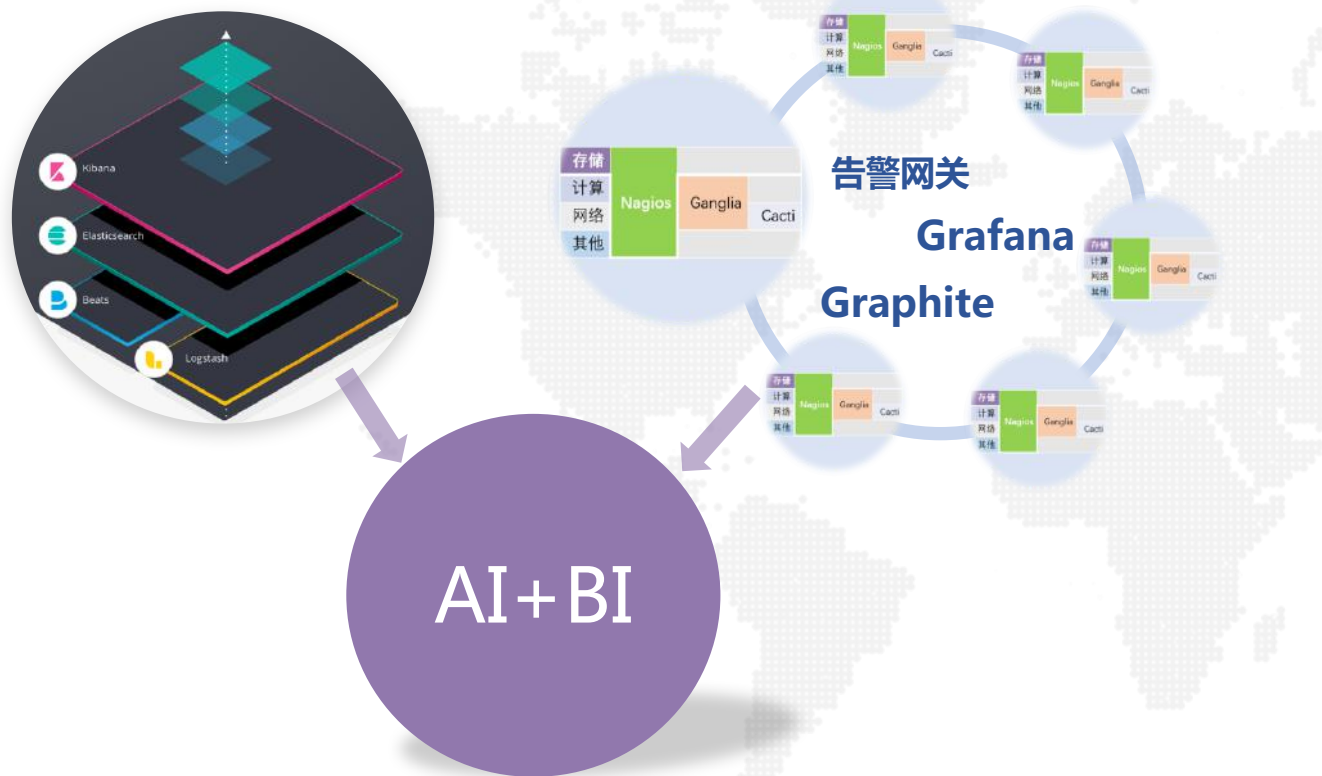
## 统一认证 :

深度定制OpenLDAP

根据多区域特性定制schema，定制认证策略，

# 华大基因高性能计算系统全球化运维

AI+BI



**AI :**

基于调度系统海量JOB日志数据（JOB特征分析与学习），全系统海量监控数据（系统效能关联分析与学习）的用户行为分析、智能运维与智能调度。

基于分布式文件系统海量数据日志，全系统监控数据的用户行为分析与存储智能HSM。

**BI :**

基于业务数据，日志数据，监控数据进行业务分析与预测，业务优化，促进企业商务智能。

# 华大基因高性能计算系统全球化运维

挑战

存储

年增量 > 100PB  
成本控制  
海量数据管理

计算

基因数据快速分析需求

# 华大基因高性能计算系统全球化运维

探索

## 存储

Lustre HSM + Quip , 对象存储  
FGPA压缩、QAT压缩、开源并行压缩算法集成  
iRODS、知识图谱

## 计算

基于日志系统业务特征分析的计算资源与调度系统优化  
基于深度学习的计算资源与调度系统优化  
FPGA硬件加速、新型软件加速  
Hadoop、Spark



# 华大基因高性能计算系统全球化运维

交流



FPGA

对象存储  
HSM

网格调度系统

Lustre 分布式存储

Hadoop , Spark

深度学习 , 知识图谱



**THANKS**