

Deployment and Monitoring for Remote Sites

Wei Zheng @IHEP-CC
zhengw@ihep.ac.cn

Outline

- Motivation
- Solutions
- Current status
- Next work
- Summary

Motivation

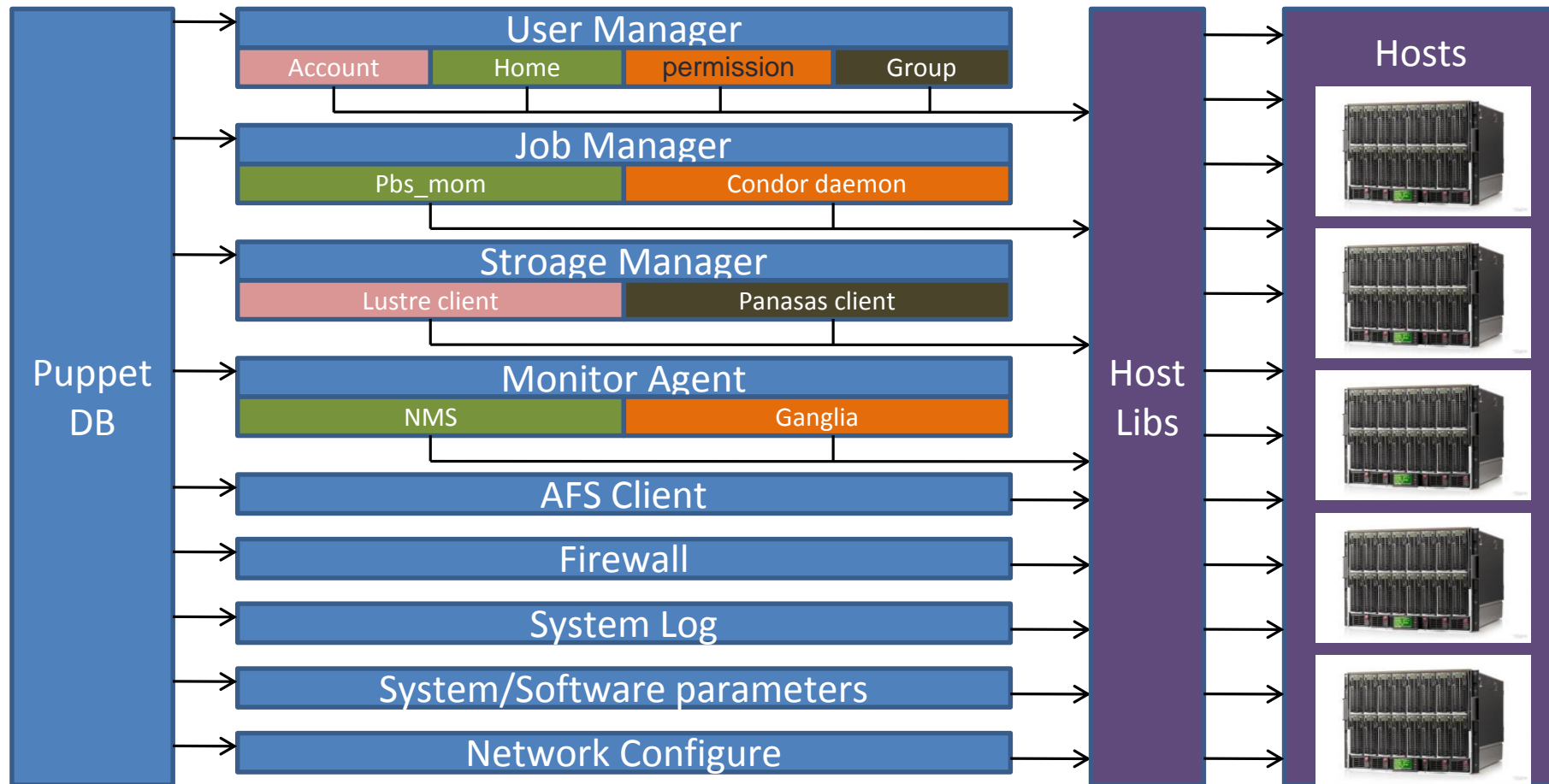
- Remote Sites of Distributed Computing
 - Jobs are dispatched to the remote sites via Dirac.
 - Stable remote sites are important for distributed computing.
- Remote sites of Distributed computing in China
 - Some experiments like LHAASO located in remote areas which face hard conditions
 - Unattended machine room
 - Small scale cluster
 - Short of man power for maintenance.
 - High security risk
 - Unstable system

Site Name	CPU Cores	Storage (TB)
Ustc	1088	1843
buaa	160	81
clas	224	37

Solutions

- IHEP provides maintenance for those remote sites
 - Software deployment and upgrade
 - Healthy status monitoring
 - Errors recovery
- Remote Site Deployment
 - Puppet & Foreman
- Remote Site Monitoring
 - Nagios & Mod-gearman

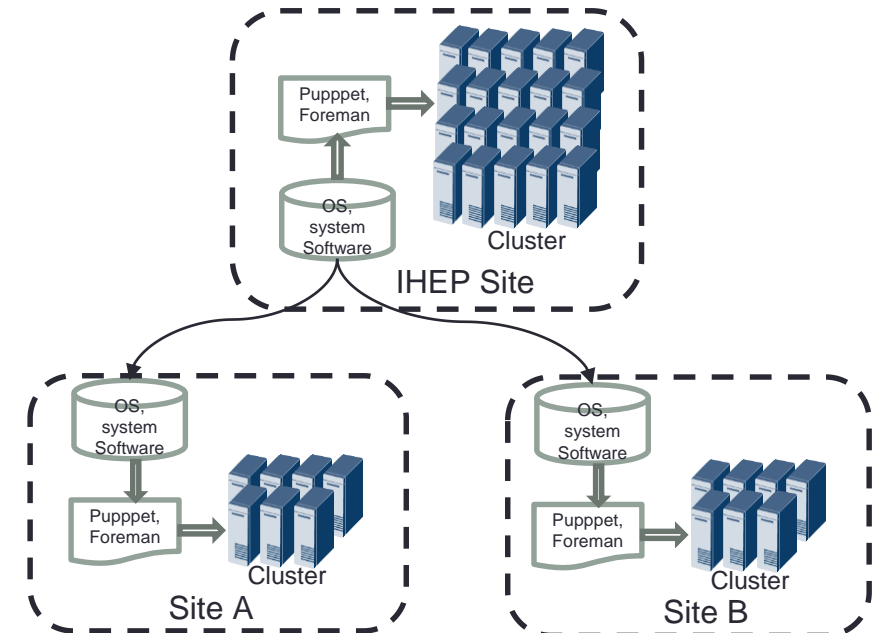
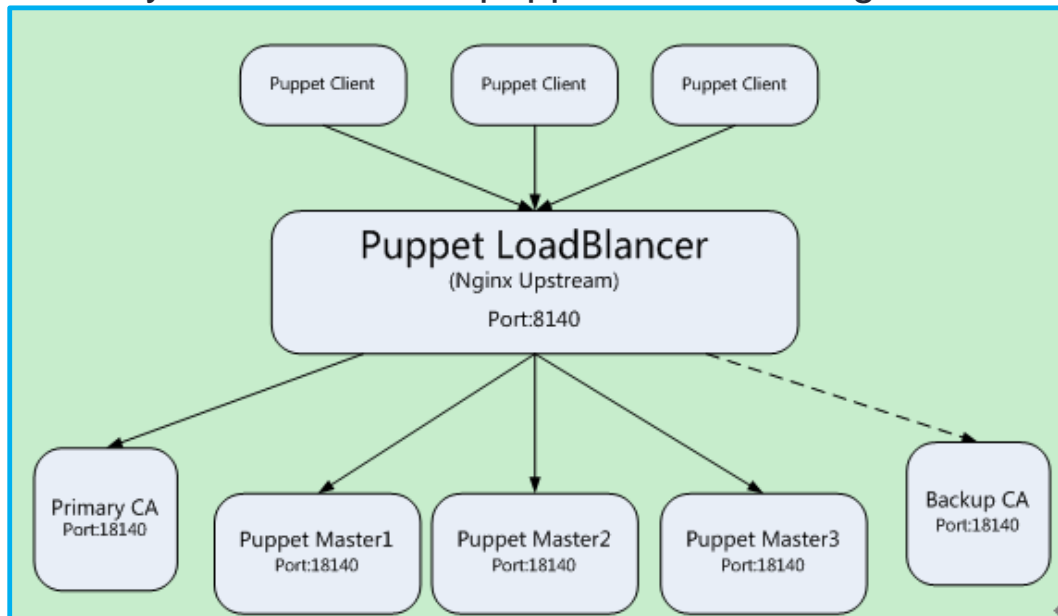
Deployment at IHEP



Distributed deployment

- Local puppet cluster:
 - 3 puppet masters + 1 puppet ca + 1 mirror host
 - Load balance within 3 puppet masters
 - Software upgrade to 2000+ hosts less than 15min
- Remote site
 - 1 puppet master

Synchronization of puppet module via git



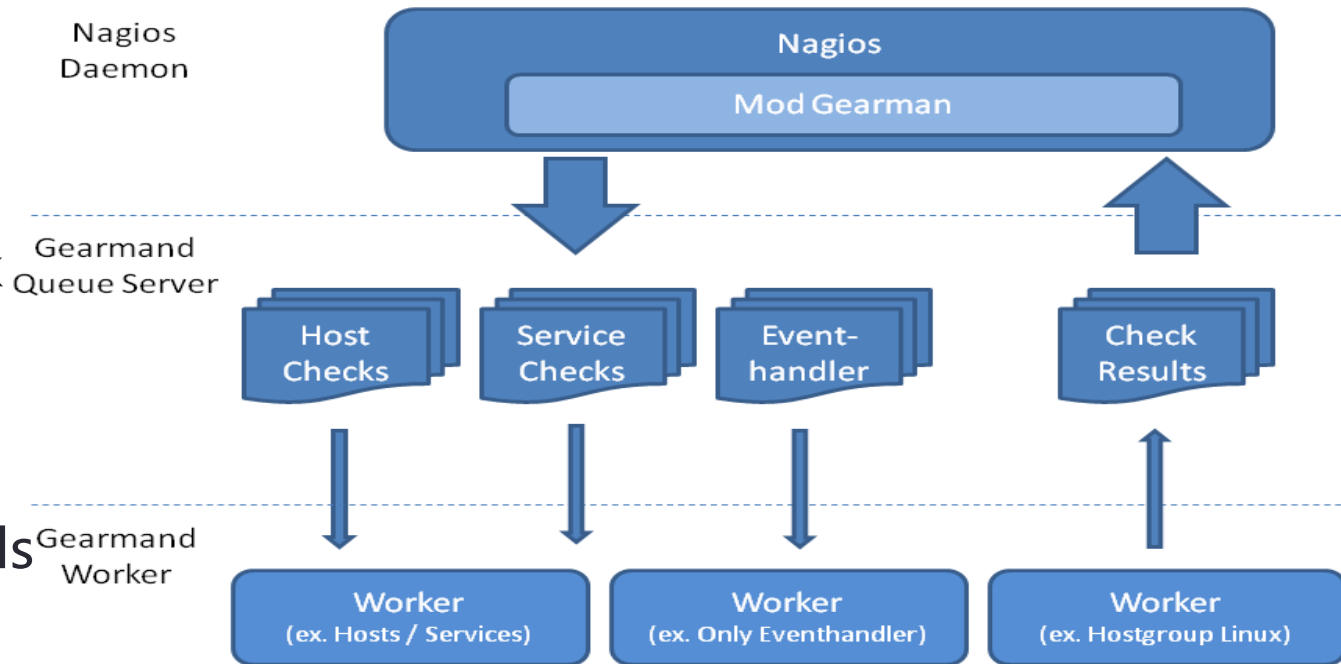
Evaluation of distributed monitoring tools

- Evaluated DNX, check_mk, Nagios+Mod_gearman
- Nagios & mod_gearman have advantage of easy configuration and high scalability

	Nagios&DNX	Check_mk	Nagios&mod_gearman
Distributed monitoring	Yes	Yes	Yes
Easy to expand	no	Yes	Yes
Load balancing	Yes	No	Yes
Performance load	No	Yes	Yes
Support proxy monitor	No	No	Yes
Check job queue	No	No	Yes
Configuration in central server	No	No	Yes

Workflow of Nagios + ModGearman

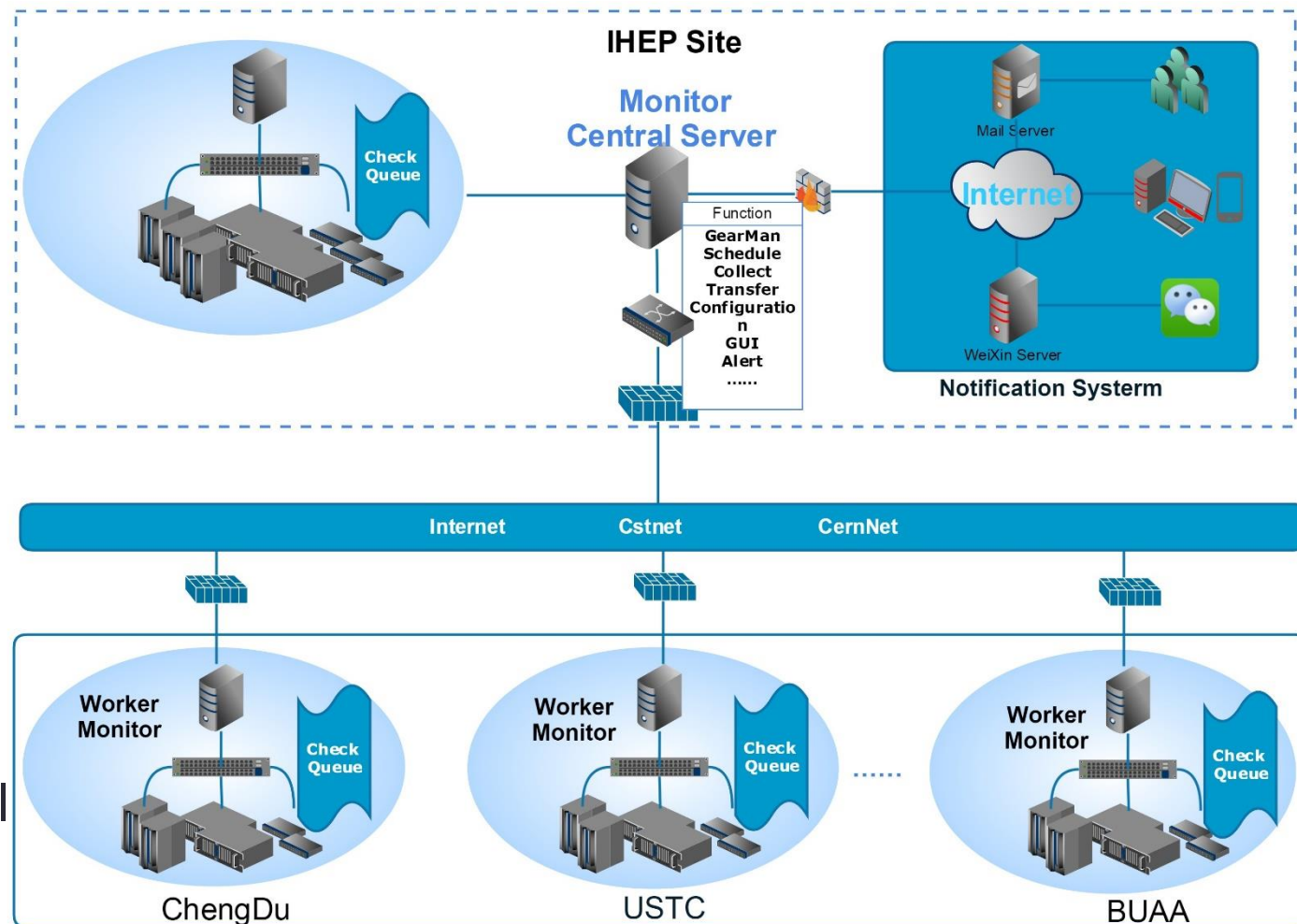
- Nagios Daemon loads a NEB (Nagios Event Broker) module
- NEB Module schedules all check jobs to the Gearmand queue
- Workers running on distributed servers execute check commands
- Mod-Gearman grabs the result job and puts back the result onto the check result list
- The Nagios reaper reads all checks from the result list and updates hosts and services



Distributed monitoring realization

Central Site: IHEP

- Nagios V4 integrate with Mod_gearman
- Receive all sites Info from the central site
- 2 worker nodes run 10 workers
 - Load balance between worker nodes
 - 13000+ services checked
- Remote Sites: Chengdu,.....
 - Reports Sites Info to central site
 - 1 worker node 5 workers at each site



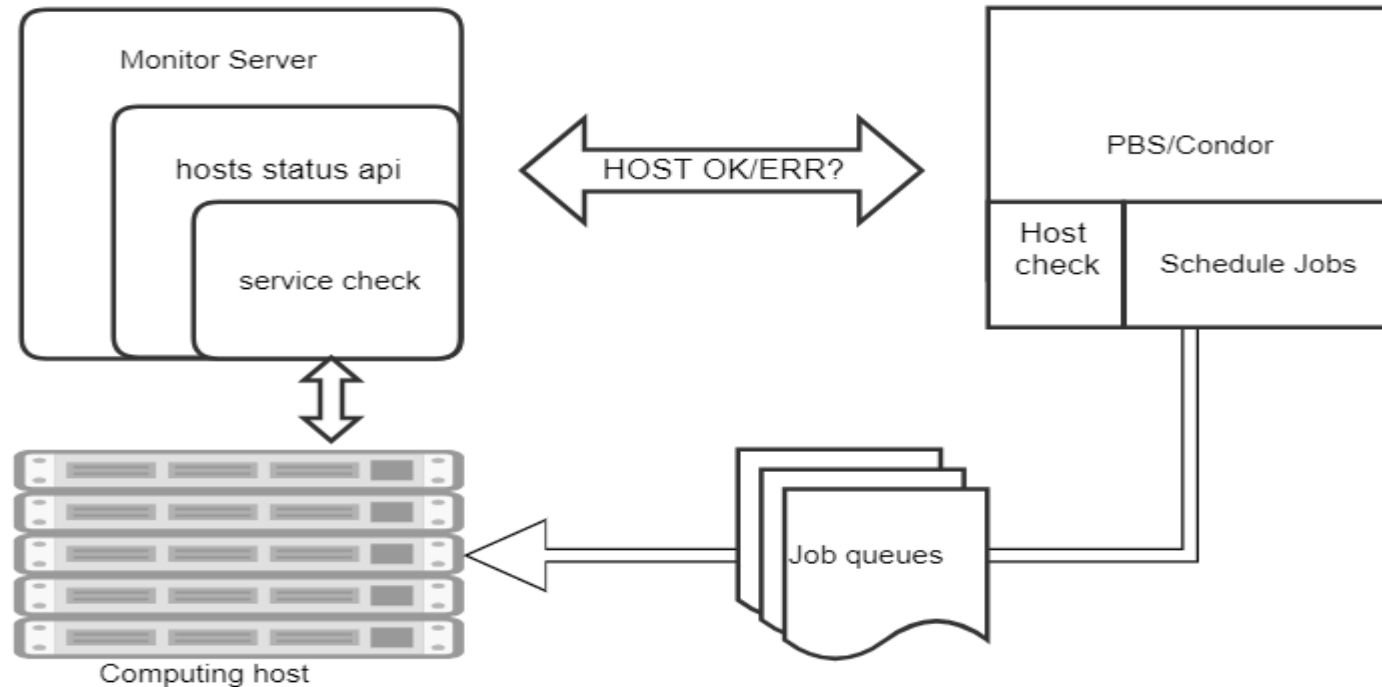
Monitor queue

```
2016-11-08 22:09:12 - localhost:4730 - v0.33
```

Queue Name	Worker Available	Jobs Waiting	Jobs Running
check_results	1	0	0
eventhandler	23	0	0
host	11	0	0
hostgroup_BUAA_Lustre_Severs	6	0	0
hostgroup_BUAA_Queue	6	0	0
hostgroup_BUAA_UI	6	0	1
hostgroup_BUAA_puppet_Severs	6	0	0
hostgroup_BeiHangWorkNode	6	0	0
hostgroup_Buaa_Monitor_Severs	6	0	0
hostgroup_ChengDu-Que	11	0	0
hostgroup_ChengDu-Schedule	11	0	0
hostgroup_ChengDu-UI	11	0	0
hostgroup_ChengDu-WorkNode	11	0	0
hostgroup_ChengDuQue	5	0	0
hostgroup_ChengDuSchedule	5	0	0
hostgroup_ChengDuUI	5	0	0
hostgroup_ChengDuWorkNode	5	0	1
hostgroup_njnu-servers	1	0	0
service	11	0	6
worker_ChengDuWoker	1	0	0
worker_NjnuWoker	1	0	0
worker_cdcompute02	1	0	0
worker_lxslc.hep.buaa.edu.cn	1	0	0

Provide monitoring API to HTCondor/Torque

- Error detected
 - host would be removed from job queue quickly
- Error recovered
 - host would be returned to the queue quickly
- Make sure the job working on properly host
- Improve the success rate of jobs



Dashboard

- Remote site monitoring has been integrated into IHEP monitoring system.

NMS

Current Network Status
Last Updated: Sat Nov 12 17:05:53 CST 2016
Updated every 90 seconds
Logged in as zhengw

Current Status
Quick Search:

Tactical Overview
Hosts
Services
Host Groups
Summary
Grid
Problems
Services (Unhandled)
Hosts (Unhandled)
Network
Outages

Monitor Analysis
Ganglia Monitor
Web Configure
Manager

Distributed Site
IHEP Site
USTC Site
BUAA Site
ChengDu Site
Cloud Site

Reports
Availability
Trends
Alerts
History
Summary
Histogram
Notifications
Event Log

System
Comments
Downtime
Process Info
Performance
Info
Scheduling
Queue

Monitor Site

Monitor Site	Host Status Totals				Service Status Totals				
	Up	Down	Unreachable	Pending	Ok	Warning	Unknown	Critical	Pending
IHEP-CC	1385	2	0	1	15614	3	8	32	0
CloudComputing: 胡床宝	207	0	0	0	1619	0	0	0	0
USTC: 郑伟	0	78	0	0	0	0	0	156	0
BUAA: 颜田	15	0	0	0	95	3	0	0	0
Chengdu: 郑伟	36	0	0	0	178	0	2	0	0

Status Summary For All Host Groups

Host Group	Host Status Summary	Service Status Summary
AMS CWS HXMT节点负责人, 闫晓飞 (AMS-Servers)	151 UP	1689 OK 1 WARNING: 1 Unhandled
AWS计算节点负责人, 系统组值班人员 (AWS-servers)	43 UP	560 OK
BES3数据库服务器 (BES3_DB_SERVER)	5 UP	5 OK
bws dbws计算节点负责人, 系统组值班人员 (BWS-Servers)	477 UP 1 DOWN: 1 Unhandled	6670 OK 2 UNKNOWN: 2 Unhandled 15 CRITICAL: 14 on Problem Hosts 1 Disabled
备份服务器-姚秋玲 (Bak-Servers)	9 UP	25 OK
BIO计算节点负责人系统组值班人员 (Bio-servers)	29 UP	290 OK
计算中心节点cac ccib map nano负责人, 系统组值班人员 (CC-Servers)	46 UP	399 OK
云计算服务器-崔涛 (Cloud-Servers)	9 UP	10 OK
DWS计算节点负责人, 系统组值班人员 (DWS-Servers)	106 UP	1378 OK
数据服务器-负责人杜国红,杨毅 (Data-Servers)	10 UP	14 OK 1 CRITICAL: 1 Unhandled
GPU负责人,文硕频6067 (GPU-Servers)	123 UP 1 DOWN: 1 Unhandled	1512 OK 3 UNKNOWN: 3 Unhandled 13 CRITICAL: 13 on Problem Hosts
存储服务器 (GRASS-Servers)	16 UP	49 OK 1 CRITICAL: 1 Unhandled
负责人:系统组值班人员 (Gluster-Servers)	13 UP	78 OK
高能所网络节点 iwn cac (IHEP-Grid)	67 UP	446 OK
江门中微子计算节点 (JNWS-Servers)	46 UP	439 OK
登录节点负责人杜国红,杨毅 (Login-Servers)	66 UP	731 OK 1 WARNING: 1 Unhandled 3 UNKNOWN: 3 Unhandled

Monitoring performance

- 5 sites include 2500+ hosts running 30000+ services
- Hosts checked ≤ 5 mins
- Services checked ≤ 5 mins

Monitor Site	Host Status Totals				Service Status Totals				
	Up	Down	Unreachable	Pending	Ok	Warning	Unknown	Critical	Pending
IHEP-CC	1324	4	0	1	17475	15	0	129	0
CloudComputing : 胡庆宝	560	0	0	0	3874	0	0	18	0
USTC : 郑伟	98	0	0	0	922	5	0	0	0
BUAA : 颜田	15	0	0	0	196	0	0	0	0
Chengdu : 郑伟	36	0	0	0	178	0	2	0	0

Hosts Actively Checked:

Time Frame	Hosts Checked
≤ 1 minute:	1134 (85.3%)
≤ 5 minutes:	1328 (99.9%)
≤ 15 minutes:	1328 (99.9%)
≤ 1 hour:	1328 (99.9%)
Since program start:	1328 (99.9%)

Metric	Min.	Max.	Average
Check Execution Time:	0.00 sec	7.50 sec	4.017 sec
Check Latency:	0.00 sec	0.93 sec	0.465 sec
Percent State Change:	0.00%	0.00%	0.00%

Wechat Notification

- A popular app integrated with messaging calling and payment function, similar to : (skype+facebook+cardless applypay)
- Warning would be sent to Wechat freely
- Become the most important means of notification



系统管理组



Regular process flows of maintenance

- Central site is IHEP
 - On-line administrators responsible for all sites regular deployment
- Remote sites are monitored by Nagios
 - Every host and service are probed
 - Errors and warnings are send to IHEP monitor site
- IHEP administrator deal with all errs
 - Login remote site and recover the error
 - Contact with remote site manager with emergency error

Current Status

- Three remote sites have joined
 - Services running at each node of remote site are monitored
 - Complete all sites checks less than 5mins
 - Physicists of USTC and BUAA are quite satisfied with the maintenance service after join

Next work

- Overall availability and reliability reports of remote sites will be provided
- More convenient deployment process will be developed
- More granular monitoring services for remote sites will be added
- New monitor to virtual machines and containers running on remote sites will be integrated
- Some small sites of universities and institutions want to join in

Necessity to join as an new site

- Monitor server of cluster can be login from outside.
- Remote site provides two servers for puppet and yum (OS source).
- Open firewall's port 5666,8139 on internal network.
 - Puppet server ports 80 and 8140
 - Yum server port 80
- A remote site contact person info (Email, mobile phone number) is necessary in case of any emergency.

Summary

- Supply site deployment service for remote site
- Remote sites monitoring services have been integrated to IHEP monitoring system
- Help remote sites running more stable
- More Chinese sites will join in in future

Thanks & Question?