



第三代 e-Science 中心核心架构



姚 舸

南京大学

人工微结构科学与技术协同创新中心

yaoge@nju.edu.cn my@yaoge123.com





eScience中心



■ 面向校内外的 数字化服务

✓ 科研

✓ 教学

✓ 管理

■ 一人全栈

服务

sci.nju.edu.cn

 云盘 为您的数据保驾护航	 协同表格 一起来填表管理信息	 超级计算 高性能科学与人工智能计算	 开源镜像 极速访问，又全又大	 私服仓库 便捷换源，码力全开
 代码托管 保留代码的点点滴滴	 LaTeX 在线编译，开箱即用	 网络测速 网速慢么？	 远程控制 安全快捷，远程直达	 文档中心 帮助都在这里

人工微结构科学与技术协同创新中心

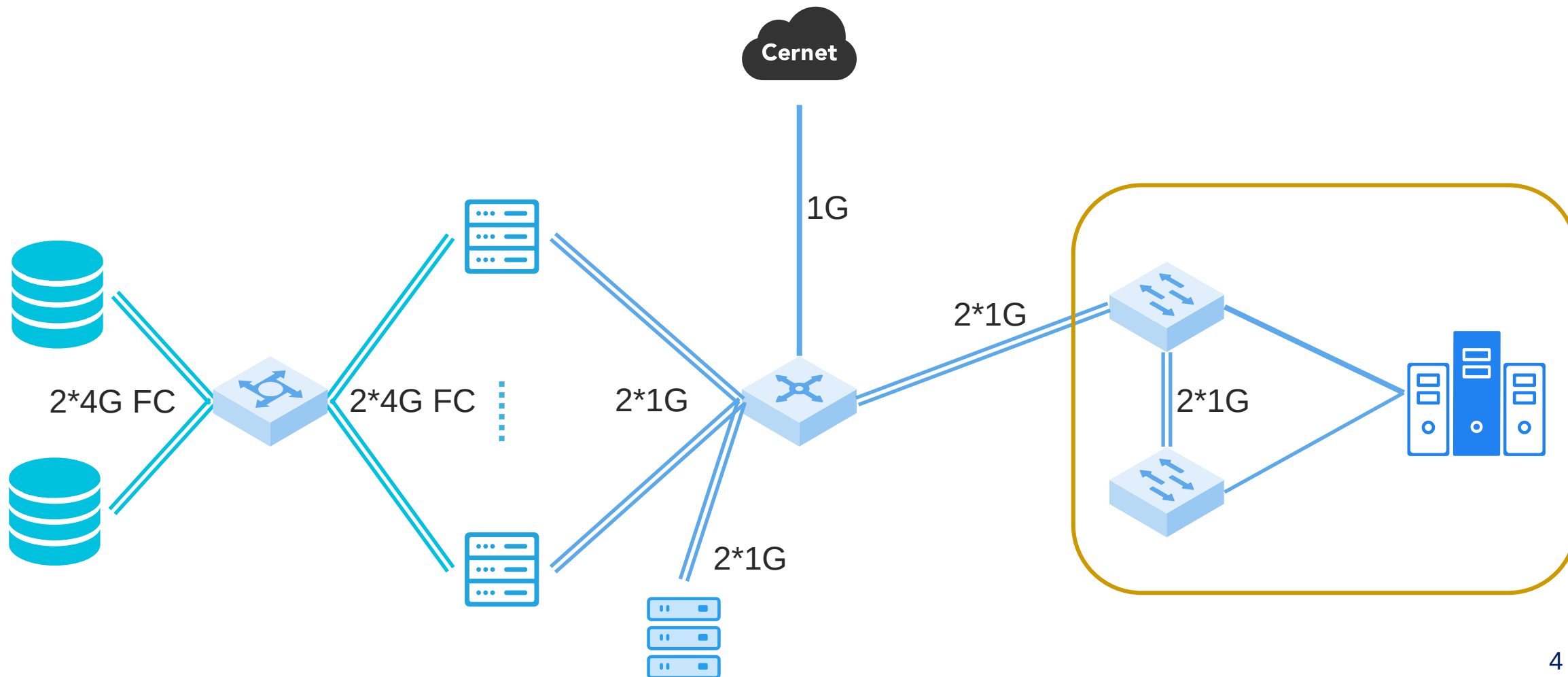


第一代



- 从零开始
 - 南京微结构国家实验室
- 主要用途
 - 管理和内部服务
- 背景
 - 迁入唐仲英楼
 - 网络自行建设

第一代





第一代



- 信息系统自行管理
- FreeBSD ZFS root
- HPC iWARP
- 网络自行管理
 - 三层上联
- 网络集中管理
 - QinQ

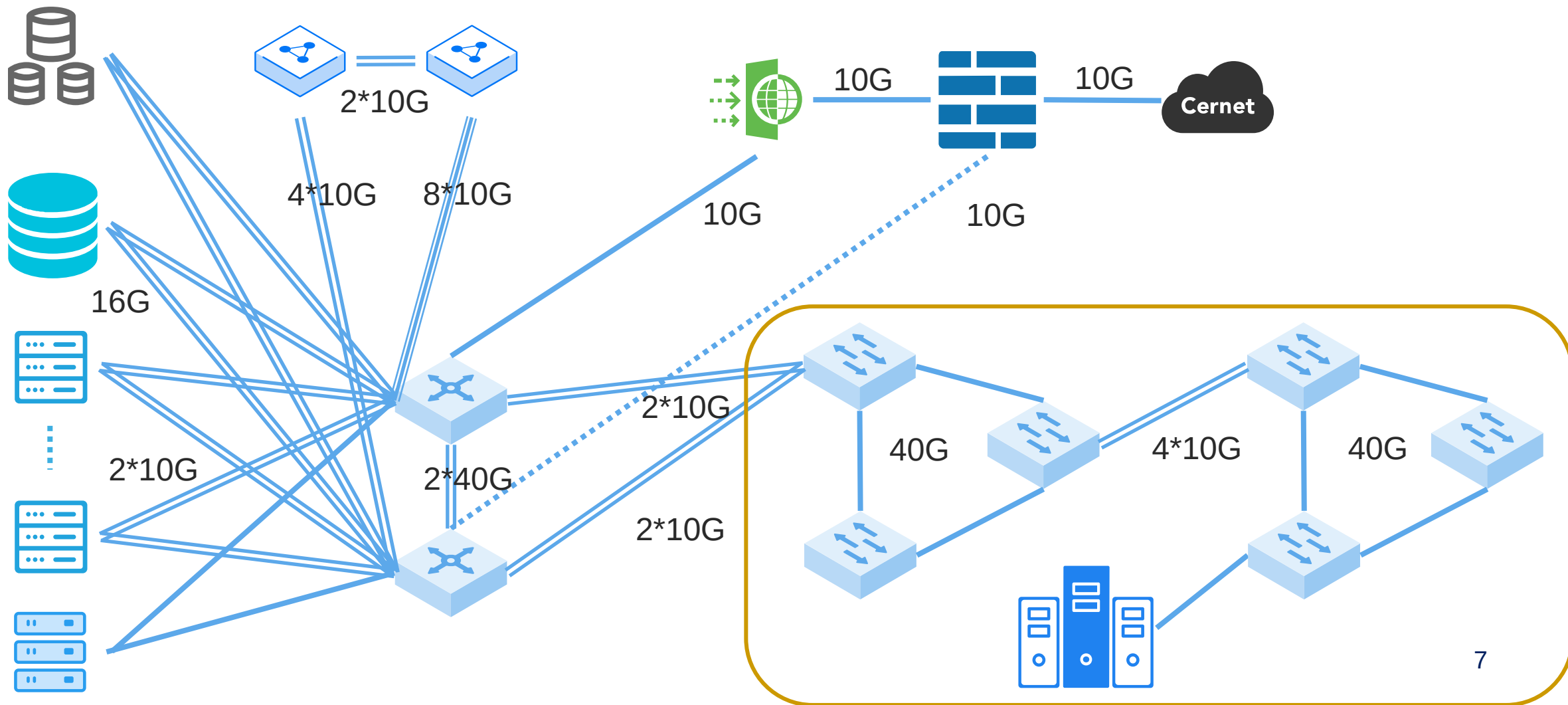


第二代



- 资源不足，更新升级
- 主要用途
 - 面校内外的广泛服务（从千人迈向万人）
- 背景
 - 教育部 财政部关于公布2014年度“2011协同创新中心”认定结果的通知
 - eScience中心

第二代





第二代



- 堆叠->STP
- FCoE
- Linux ZFS
- 独立网络
- HPC堆叠
- IB/OPA
- Lenovo Flex/AMD Rome

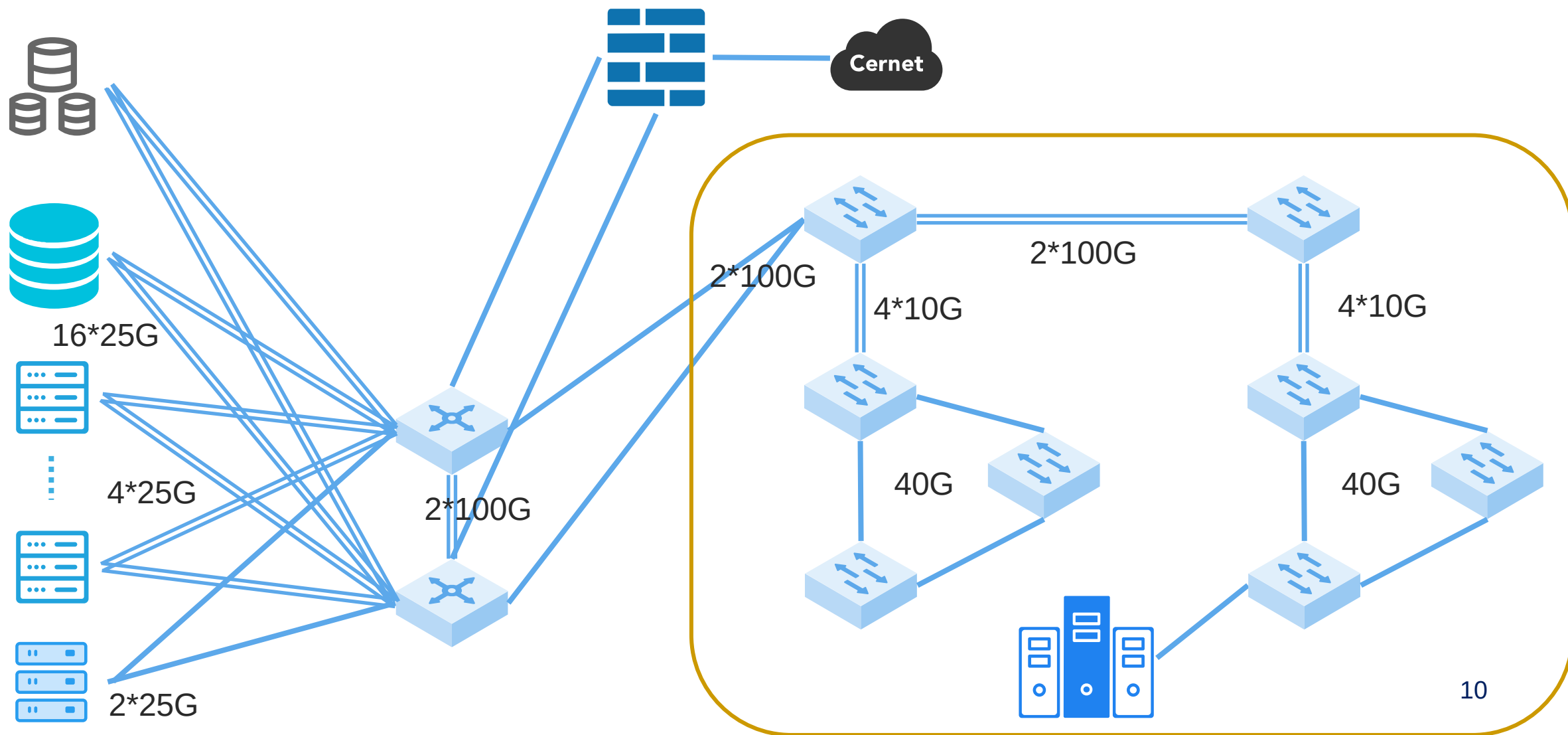


第三代

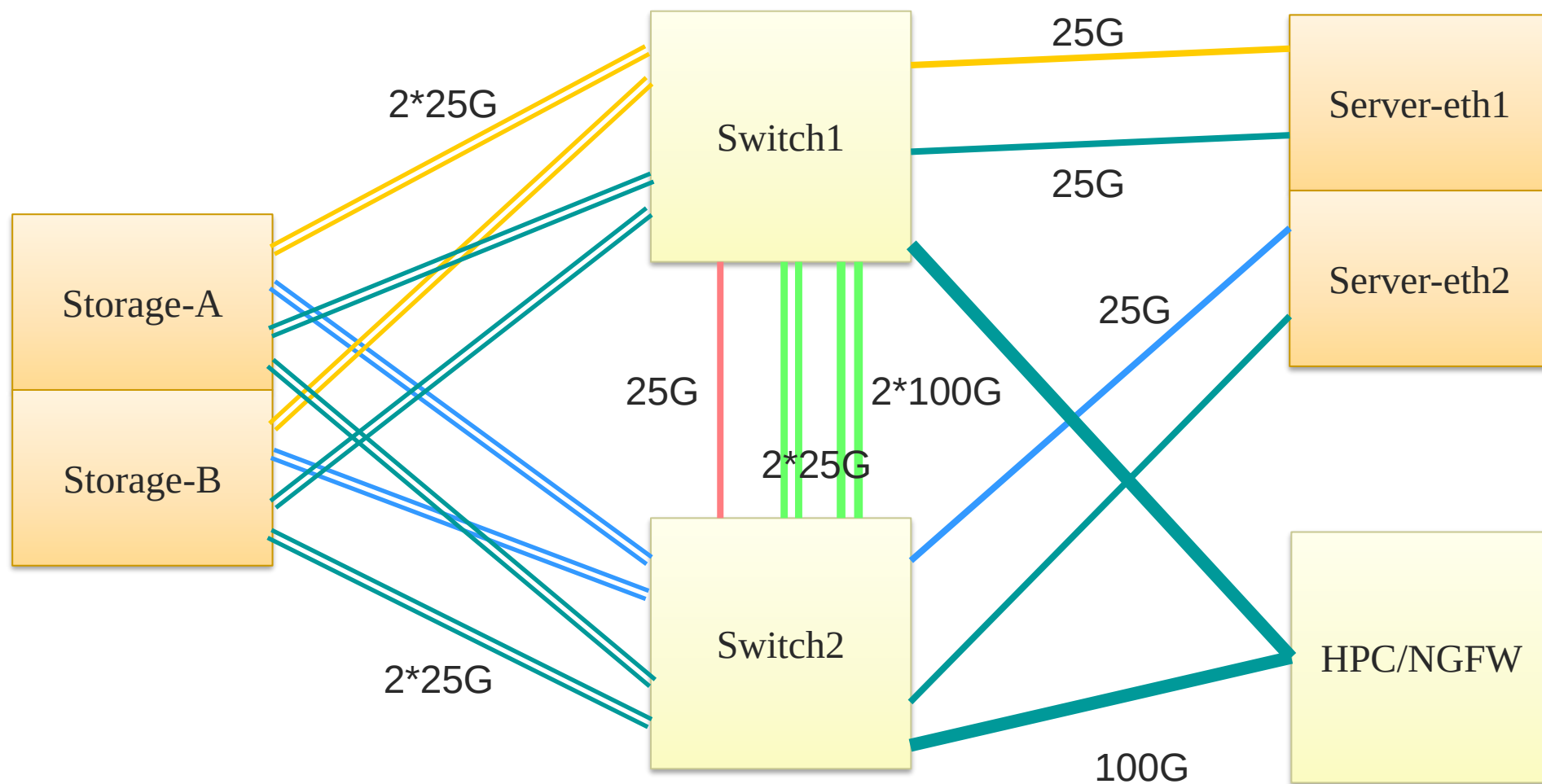


- 资源不足，更新升级
- 主要用途
 - 维持现有服务
- 背景
 - 现有设备性能严重不足
 - 核心设备EOL

第三代

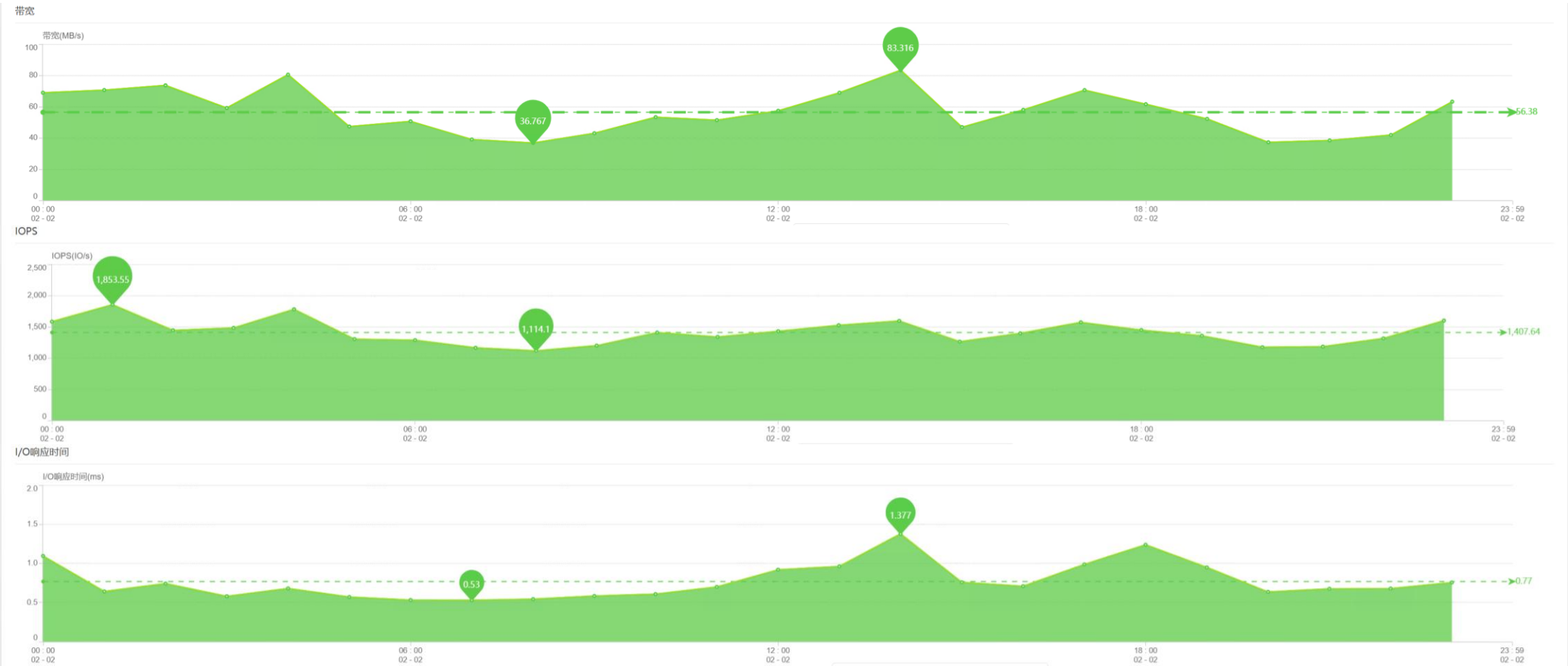


NVMe over RoCE

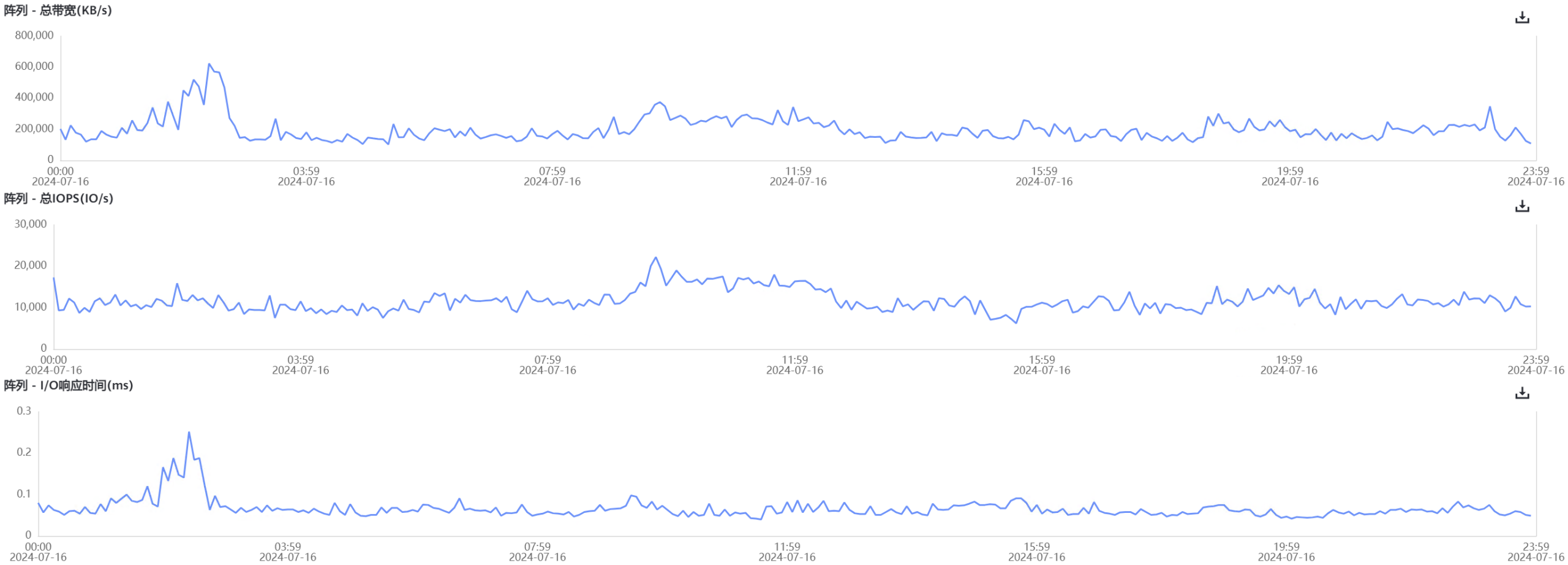




日常负载



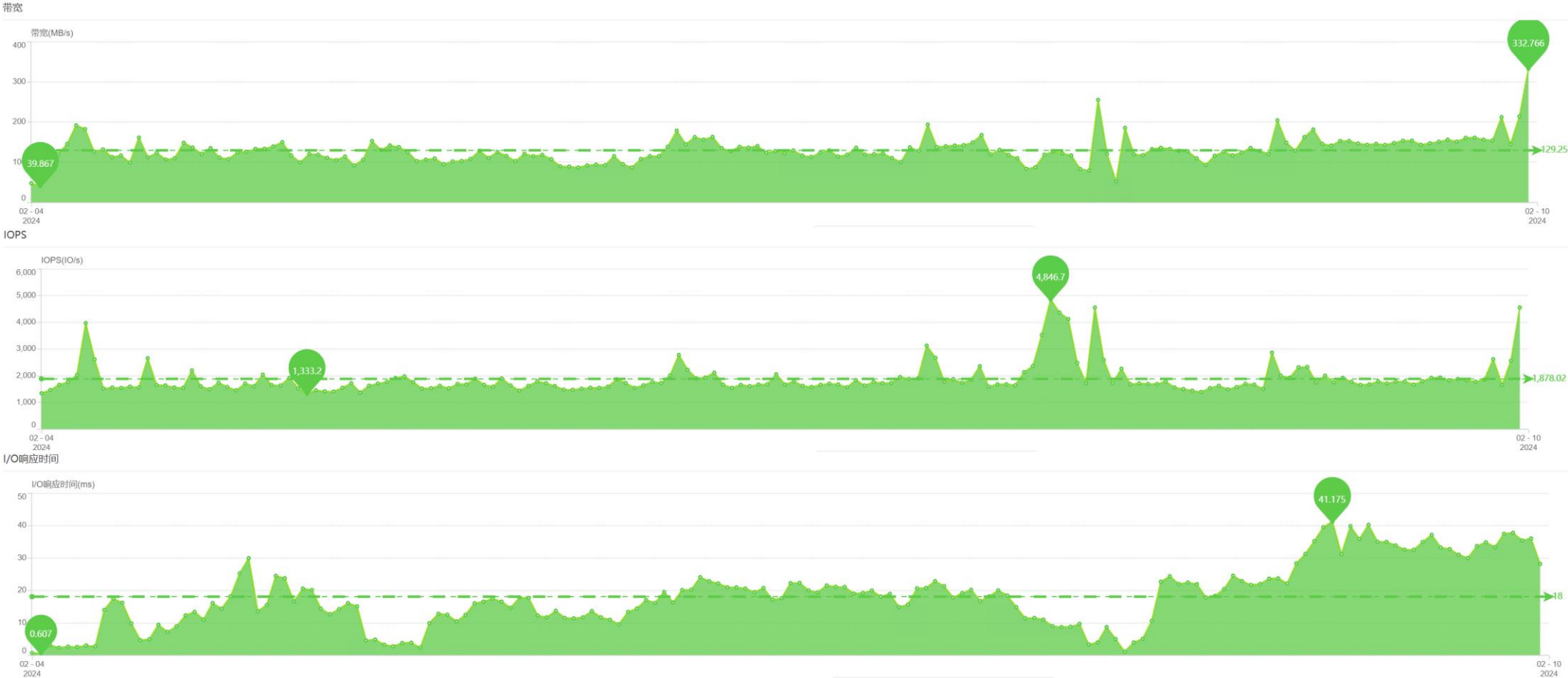
存储	带宽(MB/s)		IOPS(千)		响应时间 (ms)	
混闪	56	83	1.41	1.85	0.77	1.38
全闪	194	609	11.65	22.22	0.066	0.252



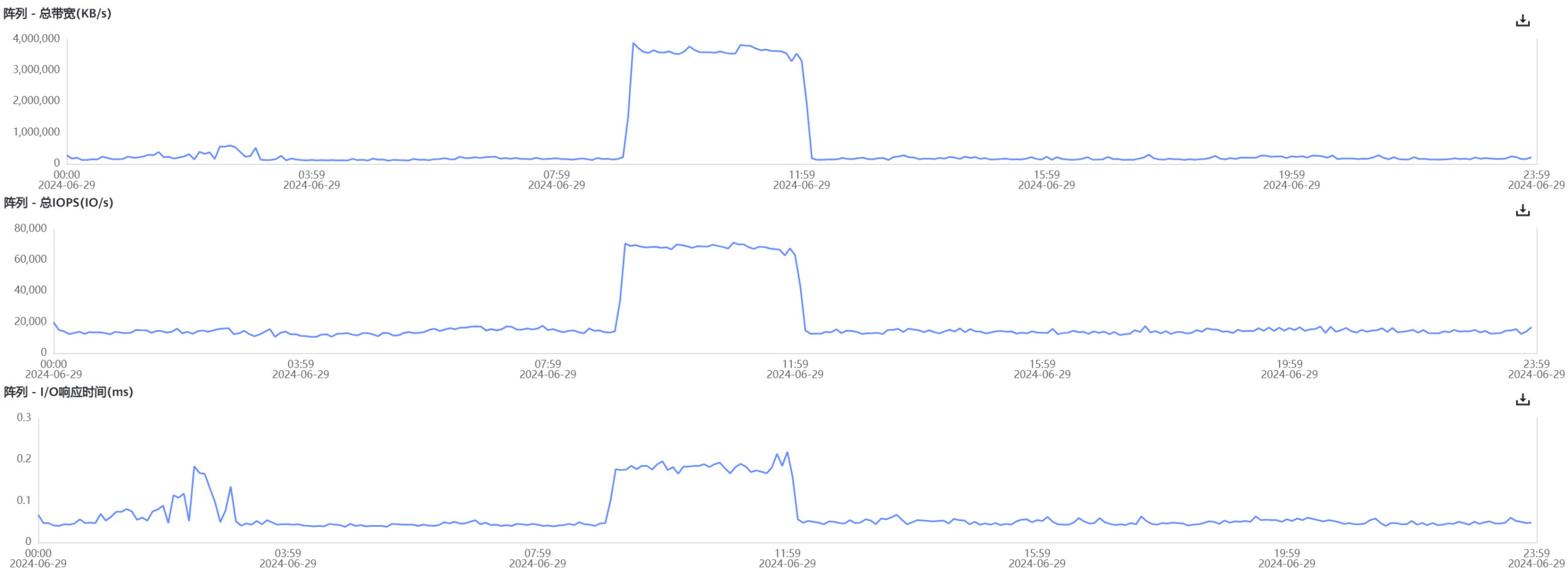
存储	带宽(MB/s)		IOPS(千)		响应时间 (ms)	
混闪	56	83	1.41	1.85	0.77	1.38
全闪	194	609	11.65	22.22	0.066	0.252



高负载



存储	带宽(MB/s)		IOPS(千)		响应时间 (ms)	
混闪	129	333	1.88	4.85	18	41
全闪	3910	4179	81	104	0.248	0.321



存储	带宽(MB/s)		IOPS(千)		响应时间 (ms)	
混闪	129	333	1.88	4.85	18	41
全闪	3910	4179	81	104	0.248	0.321

迁移后改进

■ 云盘部分文件格式增加了Cache

○ 第二十届中国研究生数学建模竞赛

华为杯研究生数模竞赛

附件数据:

1. NJU-CPOL双偏振雷达数据:

<https://box.nju.edu.cn/f/16bbb37458d3443dbf9f/?dl=1>

2. 降水格点数据:

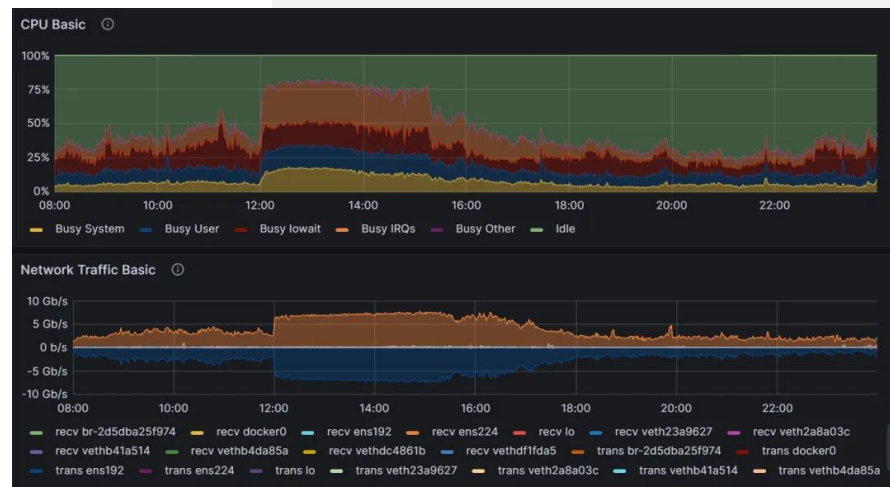
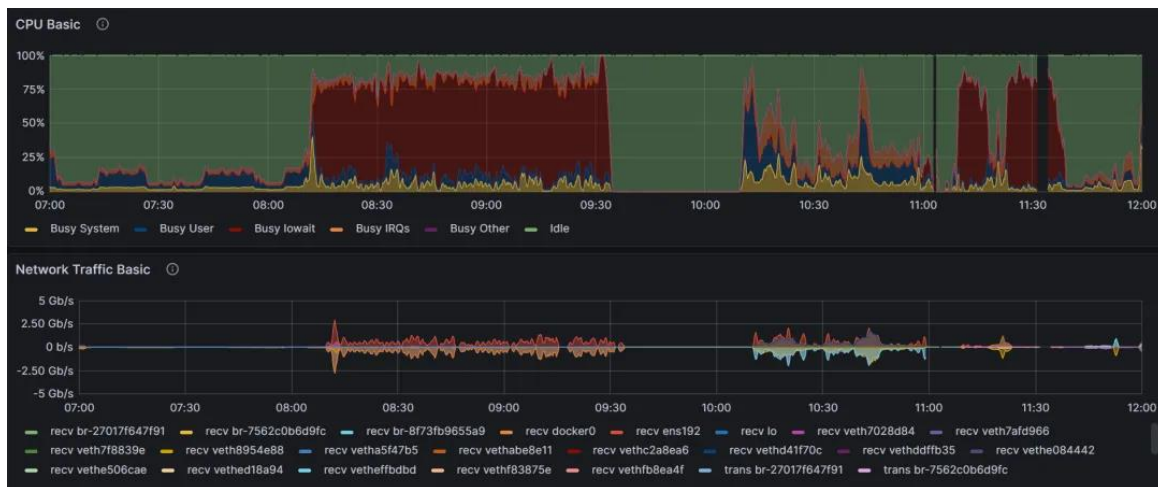
<https://box.nju.edu.cn/f/076f5aeb2ec64b87bde8/?dl=1>

估计是因为这个比赛的数据用box发放了



yaoge123 群主

今天早上的事情?





迁移后改进

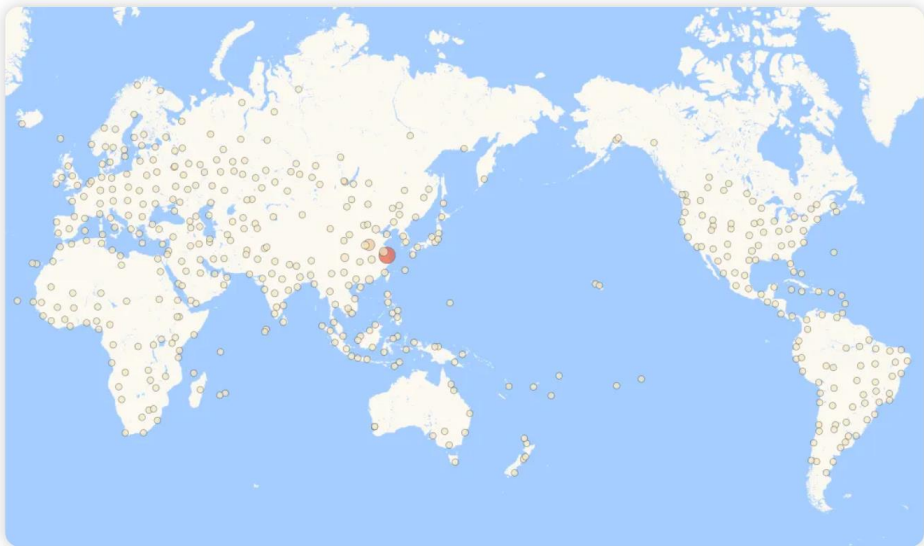


■ 关于9.22美好周五一上班就发现云盘服务异常的情况说明

^ 感谢这位群友，不然我们还蒙在鼓里.....

💰 本是开心迎周五
熟料早来急诊修

box.nju.edu.cn
南大 Box 云盘负载从 2023 年 9 月 22 日 8 点 13 分起异常升高，感受到了来自世界各地的热情：



接着陆续有校内用户发现，Box 云盘服务无法访问，其他中心服务也访问缓慢。

经过对高下载次数的链接内容的分析，确认本次异常起因是正在开展的“**华为杯**”第二届中国研究生数学建模竞赛^[1]。

本次比赛试题 F 题的数据集提供方为南大校内用户，该用户使用了南大 Box 云盘链接，**主办方未自行下载托管，而是直接使用了此链接作为下载链接且并未预先通知中心。**

由于该赛事的关注度较高、参与人员数多、参与积极性高、试题传播面广，试题发布后短时间内产生了海量并发访问。在基本无法正常服务的情况下，Box 云盘流量**暴增了100倍**，但更多的用户因无法挤入导致无法使用服务。

Box 云盘是 eScience 中心**完全本地**（换言之，所有服务器、存储数据、甚至维护人员就在南京大学鼓楼校区唐仲英楼的机房内）搭建，是主要服务校内用户科研、学习、办公的**私有云**（但允许校外访问）。

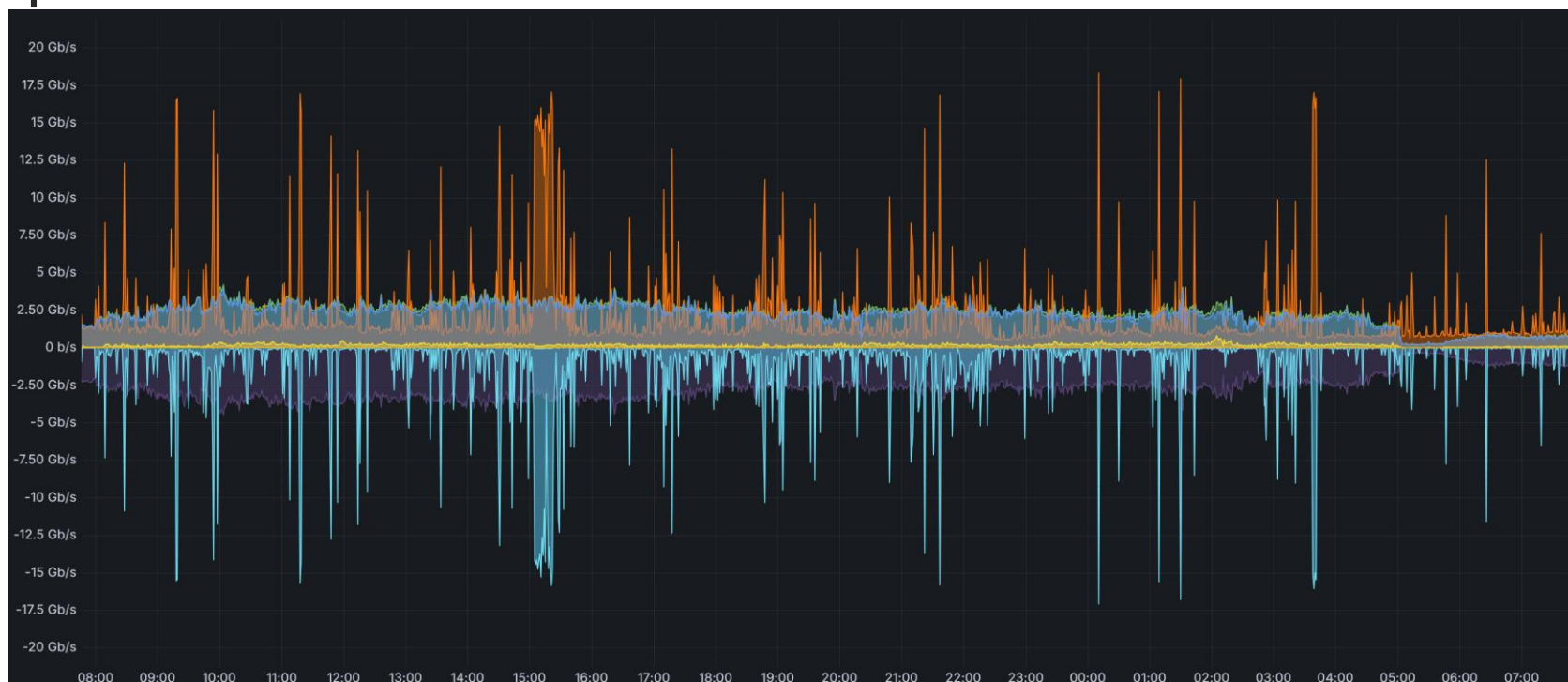
此次相对平时的访问量量级陡升，已**远超过中心用于 Box 云盘的基础架构负载能力，导致服务使用异常**。虽然“不应该把所有鸡蛋都放在一个篮子中”，但是贫穷的中心经费有限，只有一个篮子，导致中心所有业务都异常缓慢。



迁移后改进



- 开源镜像增加4T Cache
 - 降低后端ZFS压力，大幅降低rebuild时间
 - 峰值带宽接近10Gbps
 - 430种软件
 - ~285TiB





迁移后改进



■ HPC

- 切换至新堡垒机，和运维堡垒机统一
- 监控数据留存时间延长
- 抛弃Cacti，全面拥抱Grafana
- 新前端
- 百G网络



迁移后改进



- 协同表格
 - 单向备份至云盘
 - 提升Python Runner速度
- 代码托管
 - 调整存储架构
 - 提升CI/CD速度
- 网络测速
 - 满足2.5Gbps测速要求



万人应用



- 协同表格在校4.6万，离校1万
- 云盘在校2.5万，离校1.2万
- LaTeX用户1.1万，5.8万项目



双认证



- LDAP
- OAuth2

如今在各方努力之下，**协同表格成为第二个支持以此方式登入的服务，并且同时保留原先登录方式。** 接下来，中心将会继续推进Box云盘的OAuth2认证支持。

画饼完成进度：

项目	进度
Box云盘OAuth2支持	推进中😎
协同表格OAuth2支持	完成✅
手搓芯片	做梦中😭
1天扩展至48小时	平行宇宙里已经完成，参见 相关推送 👉



双认证



有了在线TeX成功集成OAuth2的经验，eScience中心也紧锣密鼓地开始规划云盘Box和协同表格的用户管理认证集成的全面升级改造，并计划在今年底、明年初**下一个大版本升级时**，推出全新用户架构的Box云盘和协同表格。届时，**将能够实现LDAP和OAuth2的双认证**，既能像原来方式登录，又可以支持校内常见的扫码登录方式。

LDAP、OAuth2、反向代理认证这些技术细节过于复杂，本文就不展开继续解释了。（有兴趣了解的师生可以给本文点个赞或转发，如果呼声高，或许可以针对这个主题开个讲座👨🎓。）

💰 为什么要等升级？深度定制不行吗？

直至19日中午，学校LDAP服务器上 `userPrincipalName` 的检索问题才最终修好，连带一些个例的历史遗留问题也一并解决。感谢信息化建设管理服务中心这段时间的辛苦维修，解决了Box云盘三年来影响最大的一个问题，为后续稳定服务提供了保障。

中心深刻反思，为什么不能做得更完美一些，能够抵御外部系统的各类问题，在这样的极端情况下仍能维持正常服务？成立以来，**eScience中心一直只有一个人，并且只有人工微结构科学与技术协同创新中心予以经费支持**。虽然理论上任何认证都可以集成，任何软件都可以定制开发。但是相关专业的师生一定明白，**过多的定制开发，必定需要更多的时间、人力和经费的投入，并会带来更多的BUG和安全漏洞**。现阶段，这的确超出了个人的精力和能力范围，可能需要等到我能够手搓芯片、一天工作48h的时候才能实现。

多年来，Seafile作为一款优秀的国产开源方案，不仅仅只有全球开源社区参与贡献，校内遇到的各类需求、问题、BUG，我们也是第一时间收集、整理并向开发团队反馈。我们的需求，都要求进入软件主线版本而非仅对南京大学定制，这**意味着不是只有南京大学可以使用这些功能，具备开放性和可持续性**。

受益于长期以来的这种开放理念，**我校用户一直以来能够享受最新Seafile Pro正式版最全面的功能**。例如，通过集成LDAP认证，云盘支持WebDAV等高级功能，这在**国内案例中是领先的**。



微信扫一扫
关注该公众号

Box云盘服务事故报告

>_ 写在最后

在经费紧缺、人手不足的情况下，eScience中心一直以最大努力去维护当前服务的正常、稳定、高质量运转。纵然，中心如果不提供Box云盘、协同表格之类的服务，就不会遭遇这样的问题——**但是，开拓创新、奋进向前，需要有“敢为、敢闯、敢干”的精神；自然，出了问题，我们也要敢担当！**

感谢广大用户一直以来的支持与关切！正是大家的支持与包容，让eScience中心在服务中进步成长。今后，eScience中心也将在成长中为广大师生提供更多更好的服务。

最后，eScience中心为给广大用户带来的不便再次致以诚挚歉意。



双测速



- 互联网: test.nju.edu.cn
- 校园网: speed.nju.edu.cn
 - 命令行测速
- ✓ 恶意滥用服务
- ✓ HTTPS
- ✓ 手机智能选网





开放资料库



■ 全校登录即可访问

- 视觉形象
- 模板
- 招生宣传
- 校园影像

公共 / 南京大学开放资料库

● 模板 7 files ● 常用 11 files ● 字体 2 files

☐	名称
☐ ☆ 	【南大宣传】近年主要招生宣传片
☐ ☆ 	【南大文创】官方文创电子资源引流
☐ ☆ 	【办公常用】免费字体
☐ ☆ 	【办公常用】南大专属模板
☐ ☆ 	【南大影像】校园文化记忆
☐ ☆ 	【南大标识】校徽、院系标识等视觉形象
☐ ☆ 	【使用须知】公共资料库内容版权与使用声明.md



服务联动



- Git x HPC x Mirrors | CI/CD自动构建
Docker镜像并用于HPC
- HPC x Table | 如何优雅地让HPC任务
完成时通知你
- HPC x Box | 在集群上使用云盘

	开源镜像、私服仓库	LaTeX	HPC
发布编译产物 皆同步发布文 发布资料库		TeX可以单向备份至Box https://doc.nju.edu.cn/books/latex/page/f033a	HPC上也可以使用Box云盘： https://doc.nju.edu.cn/books/efe93/page/352fe
			高性能计算作业信息记录、实时通知模板： https://doc.nju.edu.cn/books/18d80/page/9cfd
	替换各类源，CI/CD容器加速构建	LaTeX与Git仓库同步： https://doc.nju.edu.cn/books/latex/page/git	通过CI/CD自动构建镜像并用于HPC计算： https://doc.nju.edu.cn/books/16790/page/cicddocker

/CD容器加速			替换各类源，加速配置
---------	--	--	------------

5	LaTeX	TeX可以单向备份至Box https://doc.nju.edu.cn/books/latex/page/f033a		LaTeX与Git仓库同步： https://doc.nju.edu.cn/books/latex/page/git		
6	HPC	HPC上也可以使用Box云盘： https://doc.nju.edu.cn/books/efe93/page/352fe	高性能计算作业信息记录、实时通知模板： https://doc.nju.edu.cn/books/18d80/page/9cfd	通过CI/CD自动构建镜像并用于HPC计算： https://doc.nju.edu.cn/books/16790/page/cicddocker	替换各类源，加速配置	



南京大學

谢谢