
facebook

Scale Storage and Compute with Disaggregation

Zhilan Tan

Senior Production Engineer, Tech Lead

Table of Contents

Mark Zuckerberg

about 3 weeks ago

As of this morning, the Facebook community is now officially 2 billion people!

We're making progress connecting the world, and now let's bring the world closer together.

It's an honor to be on this journey with you.

385K 24K 11K

Challenge

The Facebook Scale

- There has never been this scale before
- Machine to machine data magnitude larger than raw data
- Maxed out all potentials
- Data processing is still seeing 10x, 100x growth

Limits of clusters

- 3+1 cluster switch forms a cluster
- Size of cluster limited by size of cluster switch
- Proprietary hardware requires special knowledge
- Impact of hardware failure is significant
- Inter cluster bandwidth is oversubscribed
- Inter cluster traffic grows

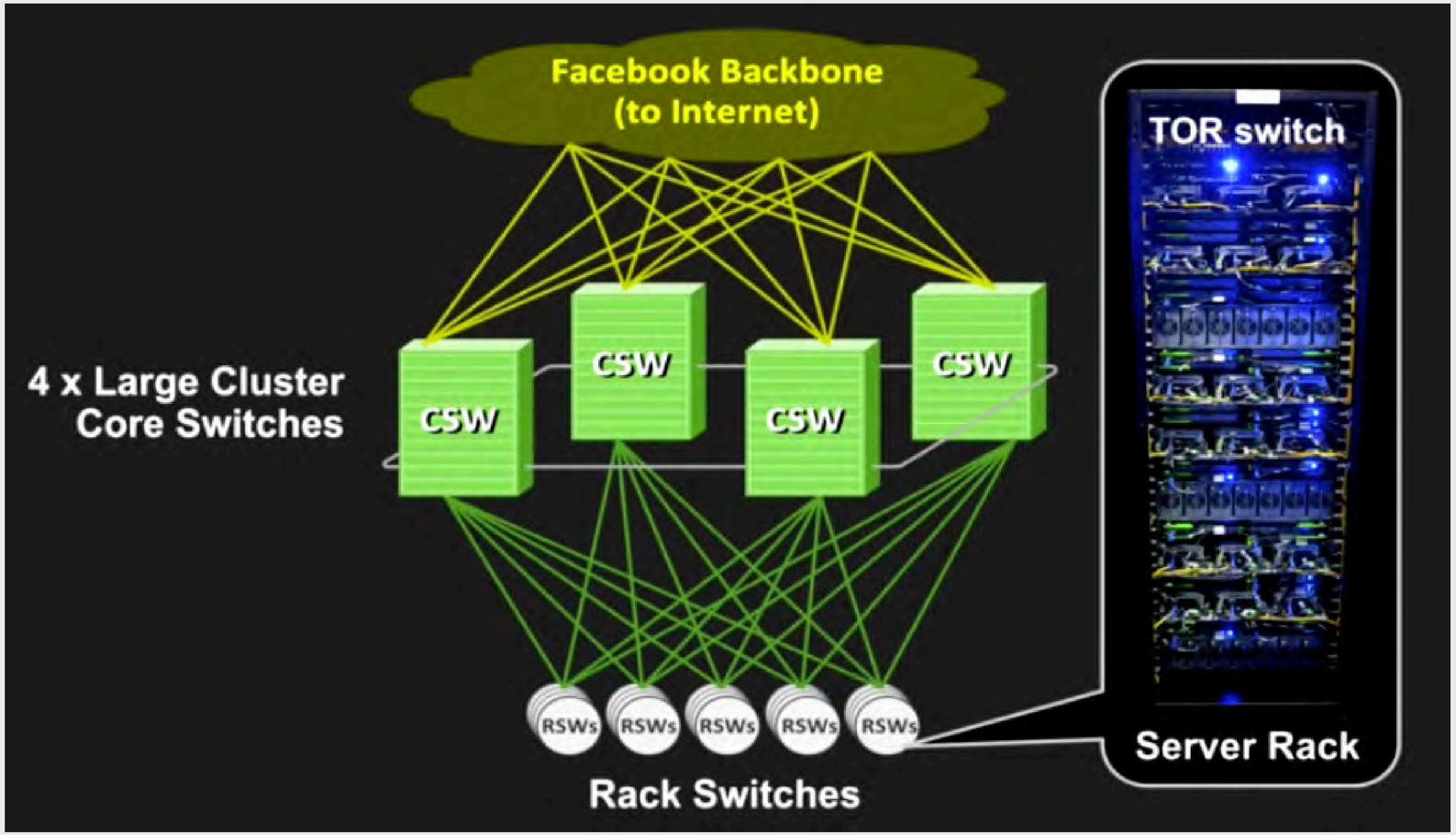
The Disaggregation Concept

Converged vs. Disaggregated

Converged	Disaggregated
Single proprietary hardware reached capacity	Commodity hardware
Hardware with software unit	Hardware and software release separately
Compute and storage colocate for data locality	Allow compute and storage scale separately

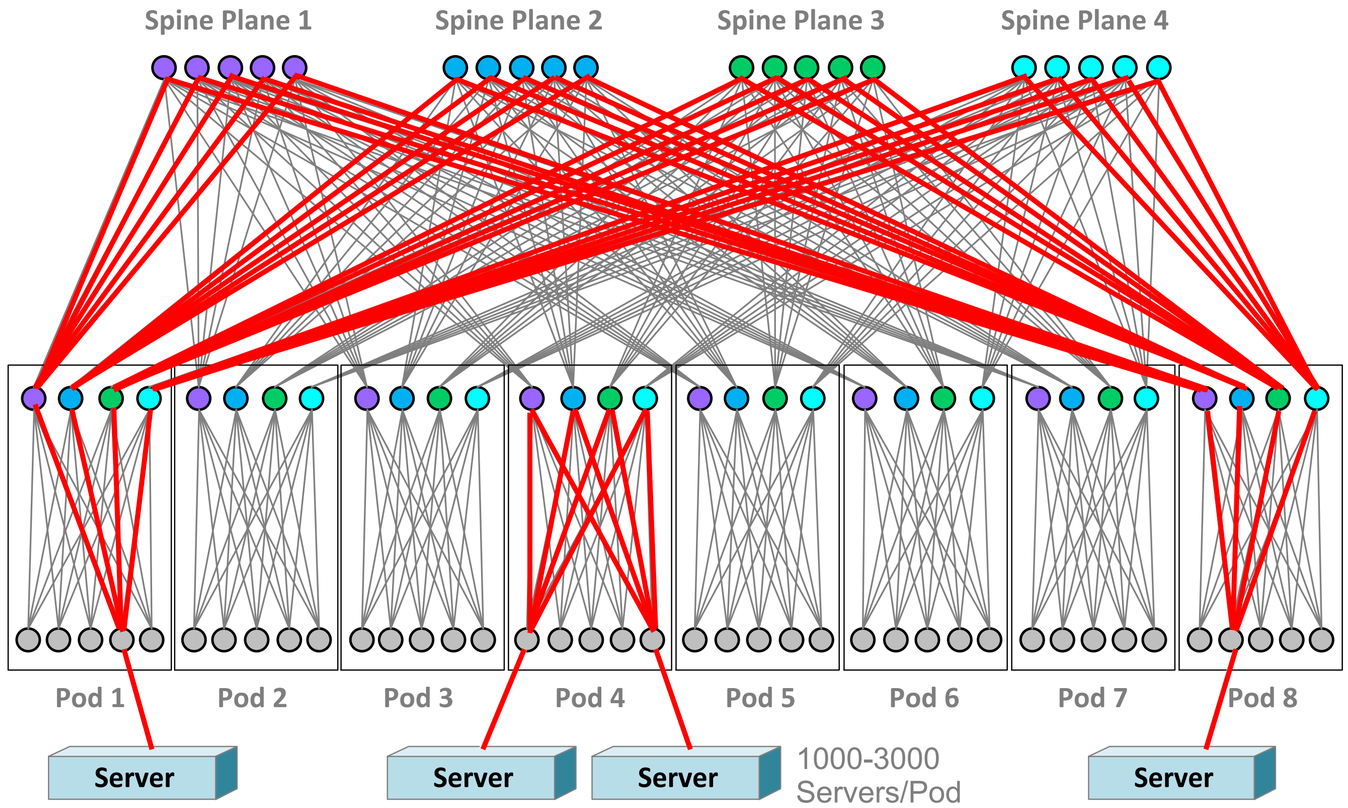
Facebook's Data Center Fabric

Making Disaggregated Compute and Storage Possible



Old Way: 4 Post Architecture

Multiple Path Between Servers



The Fabric

- Resilient to single switch failure
- Pods and planes form modular topology
- Use commodity switch
- Easy to expand capacity intra or extra fabric
- Non-oversubscribed rack to rack performance
- Expand by adding pods

With significant increase in bandwidth, reliability and scalability, disaggregate storage and compute become possible.

Disaggregate Compute and Storage

Converged vs. Disaggregated

Converged	Disaggregated
Data locality	Separate compute and storage
Unified disk/memory/cpu ratio	Different disk/memory/cpu ratio
Rely on disk performance	Network latency and bandwidth same or better than local disk
Local disk failure affect compute performance	Storage solution is resilient to disk failure

Why Warm Storage

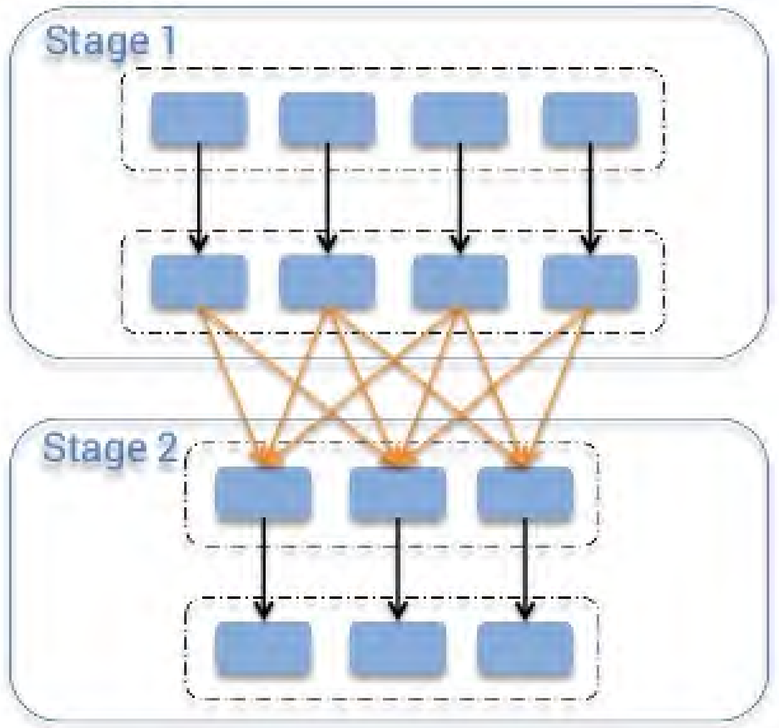
Disaggregated Storage

- Hardware SKU(StockKeepingUnit) with large disk
- High iops cause queuing
- At p99, io request can spend seconds in queue
- Disks will be slow and will fail
- Reed Solomon encoding: volume survival ratio vs. alive nodes ratio high

Spark Challenges

- Stability failure caused by disk failure
- Disk failure cause stage retry
- As job grow in size it gets worse
- Spark colo with hdfs cause io contention
- Hardware sku cause mapper memory size limit

Spark Shuffle



Spark with Warm Storage

- In production at Facebook
- On par or better than spark colocate with hdfs
- Better hardware SKU (StockKeepingUnit)
- Scale compute independently much easier
- Maintenance much easier, no hdfs draining
- 4x less failure rate

Conclusion

- Disaggregation helped us solve scalability and reliability problems
- Disaggregation is not for every situation, a smaller system may not need it


facebook

高效运维社区
GreatOPS Community

GOPS2017
Beijing

会议

- 8月18日 DevOpsDays 上海
- 全年 DevOps China 巡回沙龙
- 11月17日 DevOps金融上海

培训

- EXIN DevOps Master 认证培训
- DevOps 企业内训
- DevOps 公开课
- 互联网运维培训

咨询

- 企业DevOps 实践咨询
- 企业运维咨询

商务经理：刘静女士
电话 / 微信：13021082989
邮箱：liujing@greatops.com