



# OCP SUMMIT

March 20-21  
2018  
San Jose, CA



**OPEN**  
Compute Project



# BRYCE CANYON

## System Improvements

Madhavan Ravi/Hardware Engineer/Facebook

Yong Jiang/Storage Engineer/Facebook

Saket Karajgikar/Thermal Engineer/Facebook

CK Kho/Mechanical Engineer/Facebook



# What is Bryce Canyon?

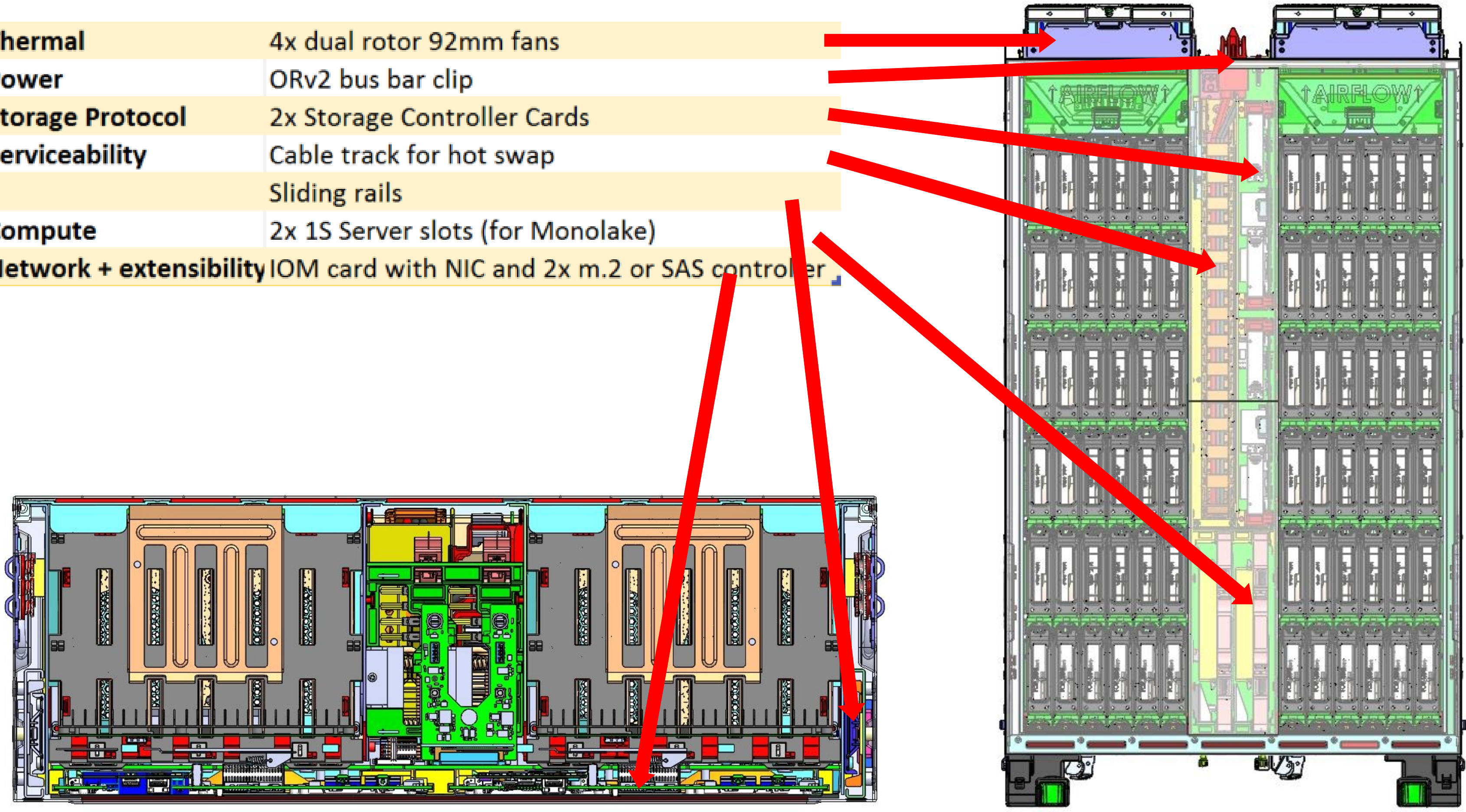
- Our latest disaggregated storage server and JBOD
- 2 storage nodes, each with 36 drives, in a single drawer
- Leverages common 1P servers (Monolake) and NICs (OCP Mezz)
- Modular and scalable design to meet current and future challenges





# System component details

<b>Thermal</b>	4x dual rotor 92mm fans
<b>Power</b>	ORv2 bus bar clip
<b>Storage Protocol</b>	2x Storage Controller Cards
<b>Serviceability</b>	Cable track for hot swap
	Sliding rails
<b>Compute</b>	2x 1S Server slots (for Monolake)
<b>Network + extensibility</b>	IOM card with NIC and 2x m.2 or SAS controller





# Why did we build Bryce Canyon?

- Dense, modular design to accommodate different SKUs with a single chassis while maintaining ease of serviceability
- Design reuse by leveraging existing microserver designs
- Improved system performance
- Efficient forced-air cooling i.e., improve CFM/W and service time
- Consistent HDD performance over all operating conditions

# Why did we build Bryce Canyon?

	Warm Storage		Cold Storage	
	Previous Generation	Bryce Canyon	Previous Generation	Bryce Canyon
Compute	Avoton 8 core	<b>Broadwell-DE 16 core</b>	Dual Haswell 12 core	Broadwell-DE 16 core
RAM per Compute	32GB DDR3	<b>64GB DDR4</b>	128GB DDR3	128GB DDR4
HDD per Compute	30	<b>36</b>	240	216
HDDs per Rack	450	<b>576</b>	480	<b>648</b>
SSD Slots (M.2) per Compute	1 x M.2 SATA	<b>2 x M.2 NVMe</b>	0	0
Max Network BW per Compute	10Gbps	<b>50Gbps</b>	10Gbps	<b>50Gbps</b>

The Warm Storage version of Bryce Canyon provides ~4x compute, 2x DRAM, consumes 30% less power / HDD, and helps achieve ~50% reduction in CFM/W



# System Improvements

Consistency in HDD performance

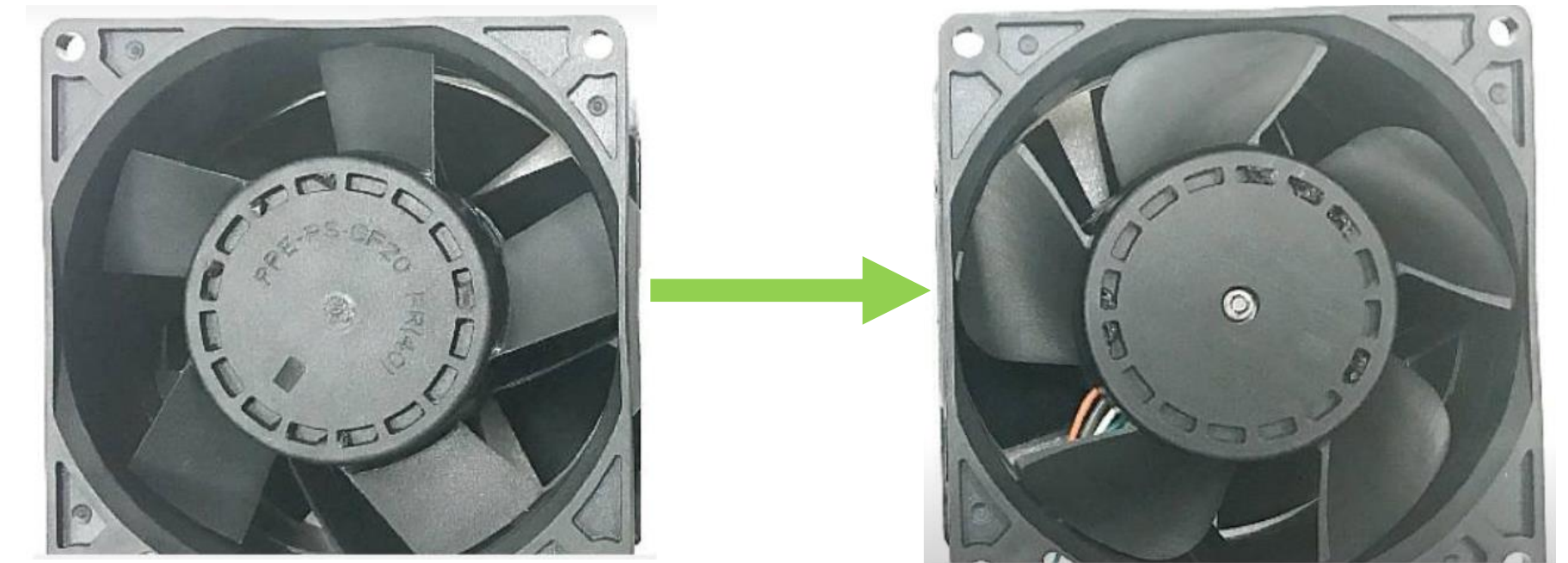
# HDD performance goals

- Sustained operating conditions
  - Goal: <5% maximum performance degradation
- Non-sustained operating conditions
  - Goal: <10% maximum & <7% average performance degradation



# Updates from acoustic/rotational vibration testing

- Observation: Higher noise levels in acoustic frequency ranges due to the fans
- Fixes:
  - Modified fan blade angle
  - Added a honeycomb layer
  - New wire-frame finger guard design



Old Fan blade

New Fan blade



Honeycomb acts as acoustic attenuator



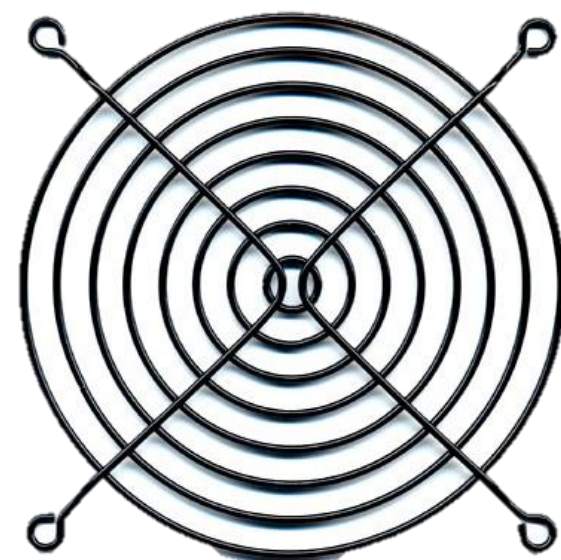
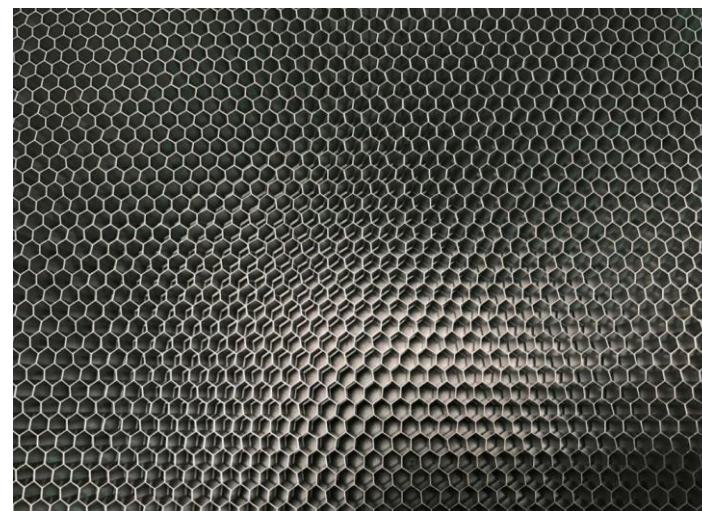
Smaller stamped metal vents



Larger wireframe guard vents



# Updates from acoustic/rotational vibration testing



Improved HDD  
Performance

	HDD Bandwidth Change in Back Row (100% PWM in Fan Speed)	
	Old Fan	New Fan + Honey Comb + New finger guard
Slot 0	-64.4%	-0.2%
Slot 1	-71.6%	-0.4%
Slot 2	-79.0%	-0.3%
Slot 3	-44.7%	-0.3%
Slot 4	-53.2%	-0.8%
Slot 5	-39.2%	-0.6%
Slot 6	-50.4%	-0.6%
Slot 7	-67.3%	-0.5%
Slot 8	-90.9%	-2.3%
Slot 9	-75.2%	-0.5%
Slot 10	-99.7%	-0.7%
Slot 11	-64.1%	-0.3%

Huge improvement in performance for the last row of HDDs at 100% PWM



# System Improvements

Thermal



# Summary of improvements

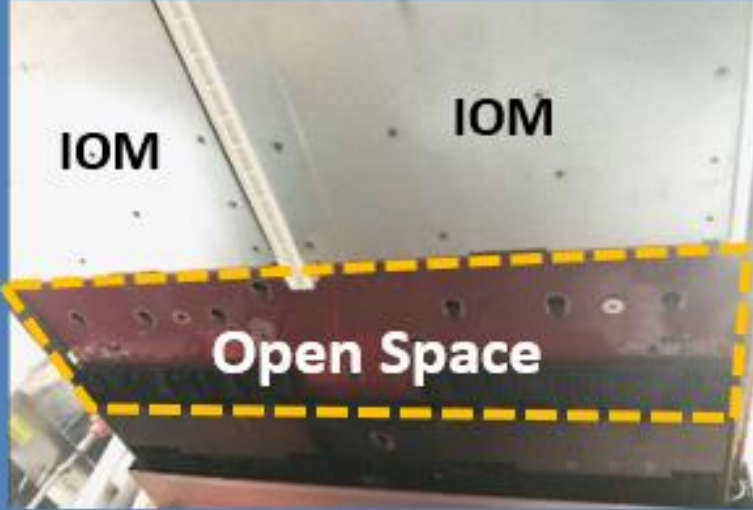
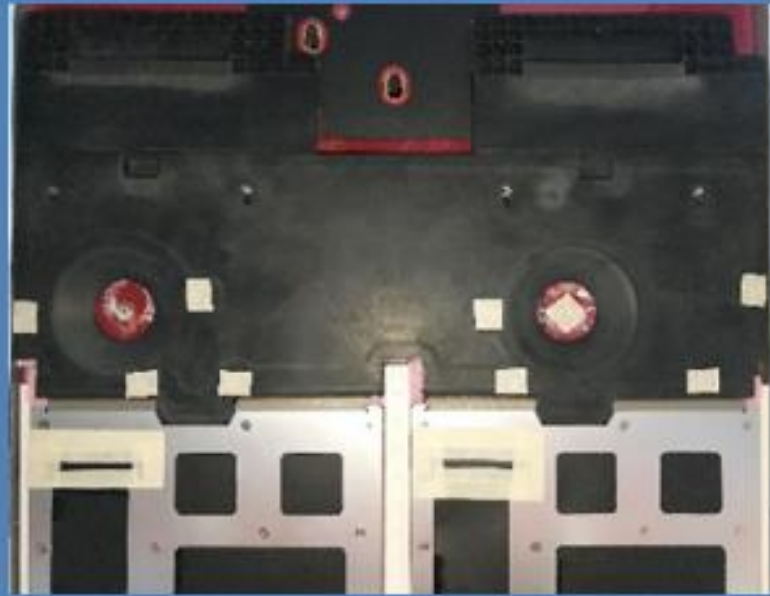

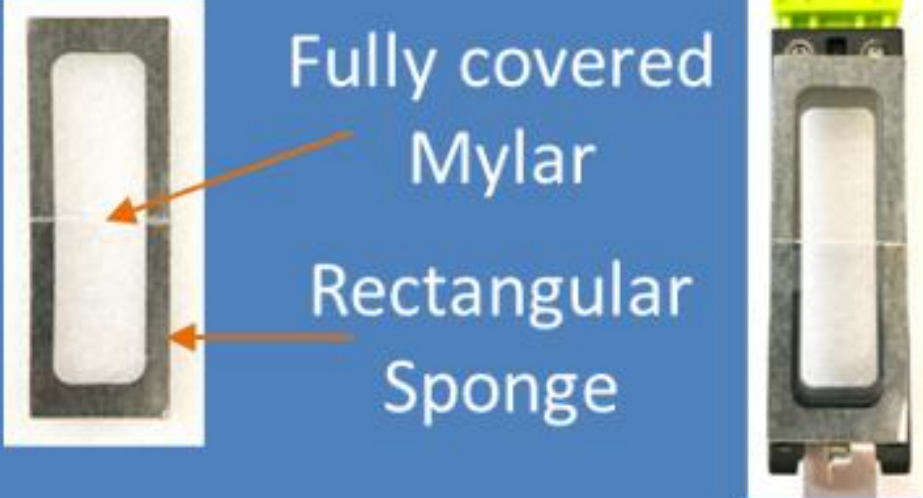
- Minor system-wide changes to improve the service time window
- Lowered CFM/W by ~5%, compared to what we started with
- Fan modifications to lower acoustic vibration intensity on HDDs

# Service time window requirements

- **What is serviceable on a Bryce Canyon?**
  - Online HDD(s) swap
  - Offline swap of the DIMM(s), server, IOM, SCC, m.2 or NIC, while the other server instance in the chassis is operational
- **System design goals**
  - System shall operate normally with acceptable performance degradations when the drawer is fully pulled out
  - System should be designed to support appropriate service time windows at all intended temperature and altitude ranges



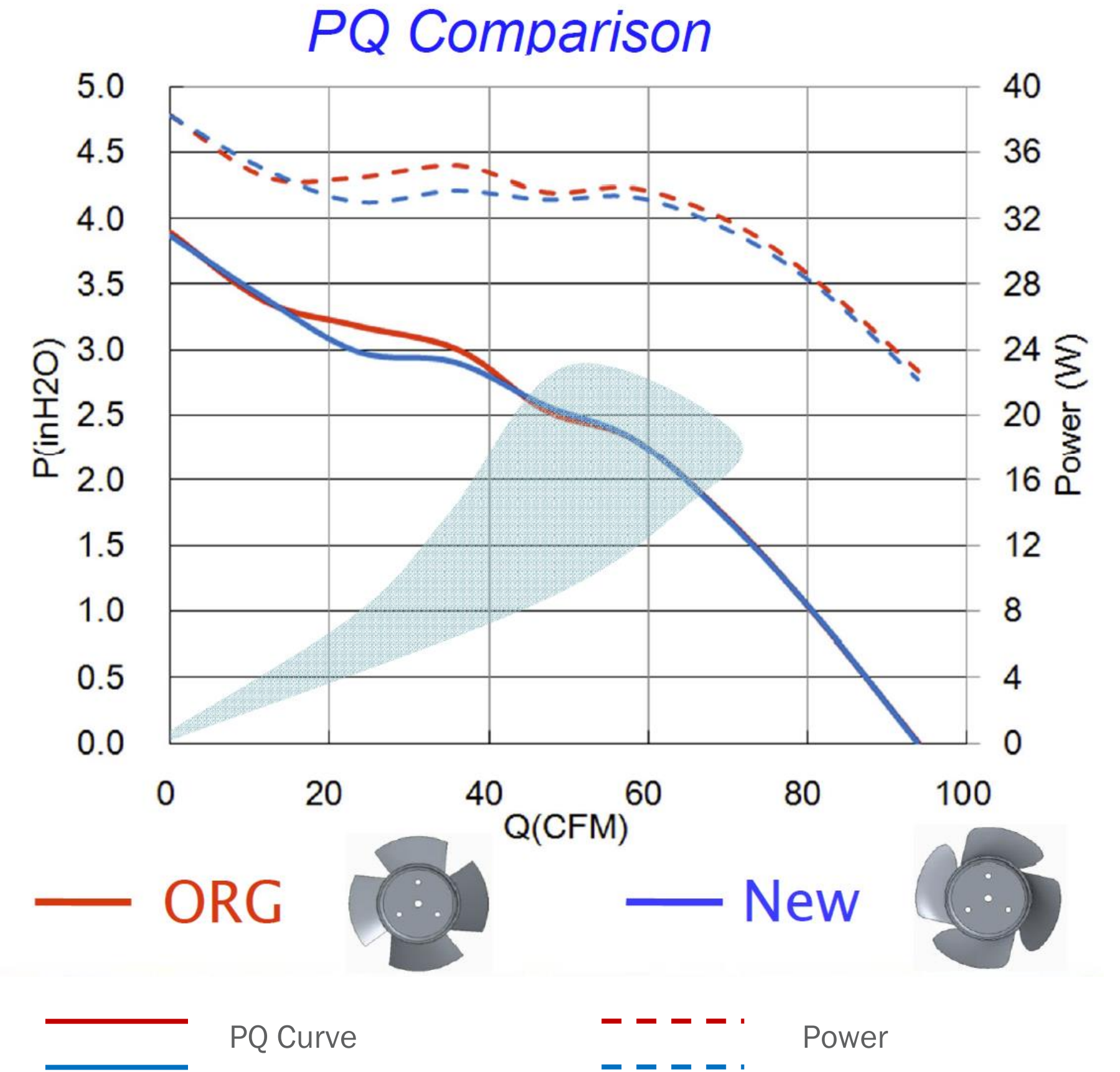
# Changes to improve service time window

	Original	#1	#2	#3
<b>Thermal Solution</b>	without any solution	A hole on IOM chassis below NIC heatsink	A hole on IOM chassis below NIC heatsink	A hole on IOM chassis below NIC heatsink
	without any solution	[Reworked CNC DPB Cover] (with gasket to seal up gap between cover and IOM chassis) <i>(Reduce the height of center ribs)</i>	[CNC DPB Cover] (with gasket to seal up gap between cover and IOM chassis)	[Reworked CNC DPB Cover] (with gasket to seal up gap between cover and IOM chassis) <i>(Reduce the height of center ribs)</i>
	without any solution		[New HDD sponge] Remove the original U-shape sponge and add a new rectangular sponge on the bottom side of HDD latch	[New HDD sponge] Remove the original U-shape sponge and add a new rectangular sponge on the bottom side of HDD latch + Mylar sheet under the sponge
				
<b>Service Time</b>	<b>193s (3.2 minutes)</b>	<b>309s (5.2 minutes)</b>	<b>508s (8.5 minutes)</b>	<b>(&gt;20 minutes)</b>



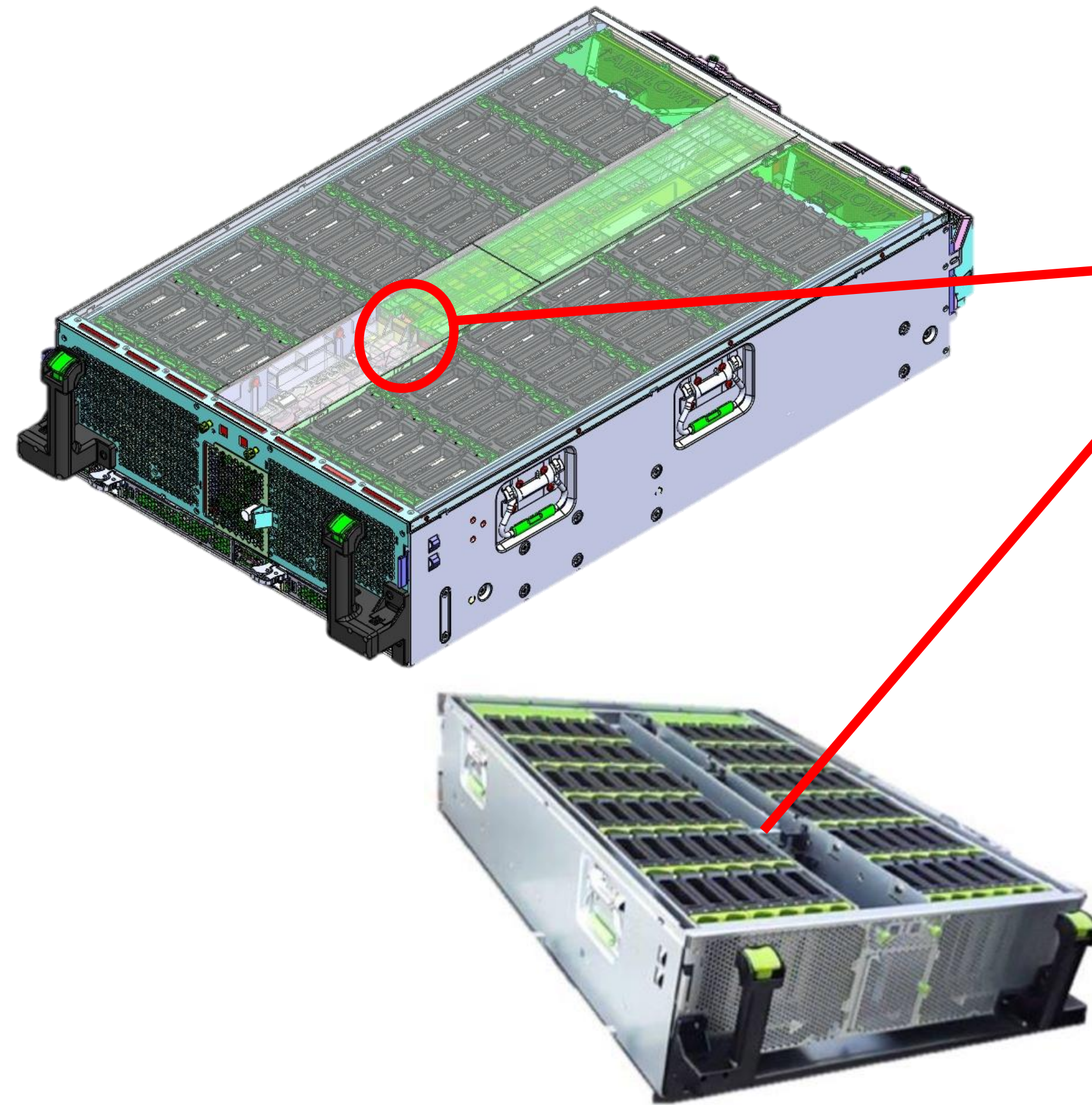
# Changes to fan blade design

- Modified the fan blade design → smoothen the blade tips
- Modified the impeller position
- Noise lowered by 3.7 dB(A) which lowered the HDD performance degradation due to acoustic vibrations
- No change in fan performance



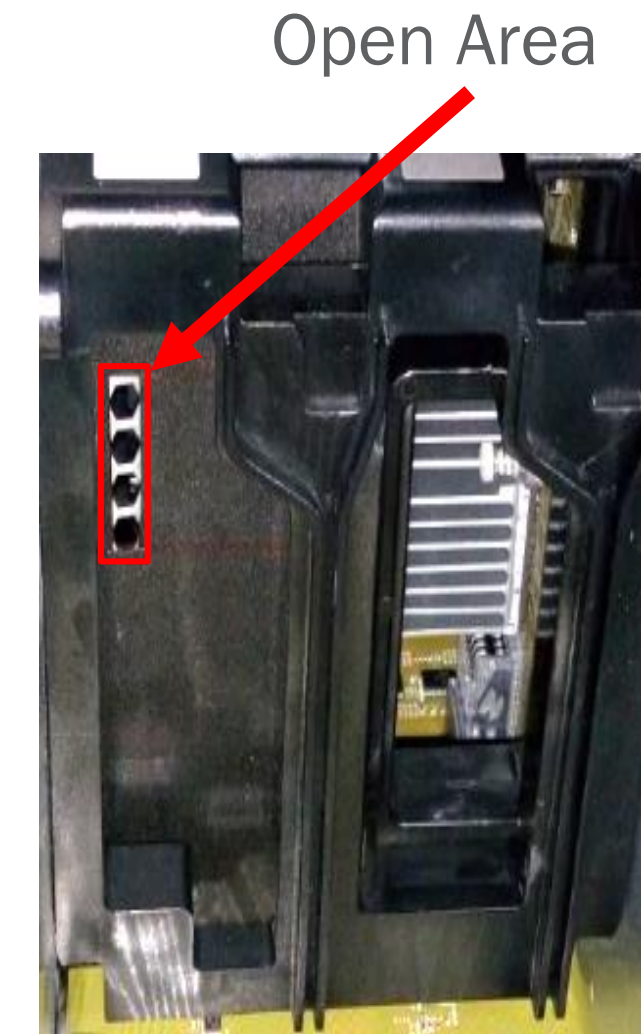


# Changes to server chamber wall



0.129 CFM/W  
@30 °C

Reduction in  
open area



0.122 CFM/W  
@30 °C

Lower CFM/W  
under identical  
conditions



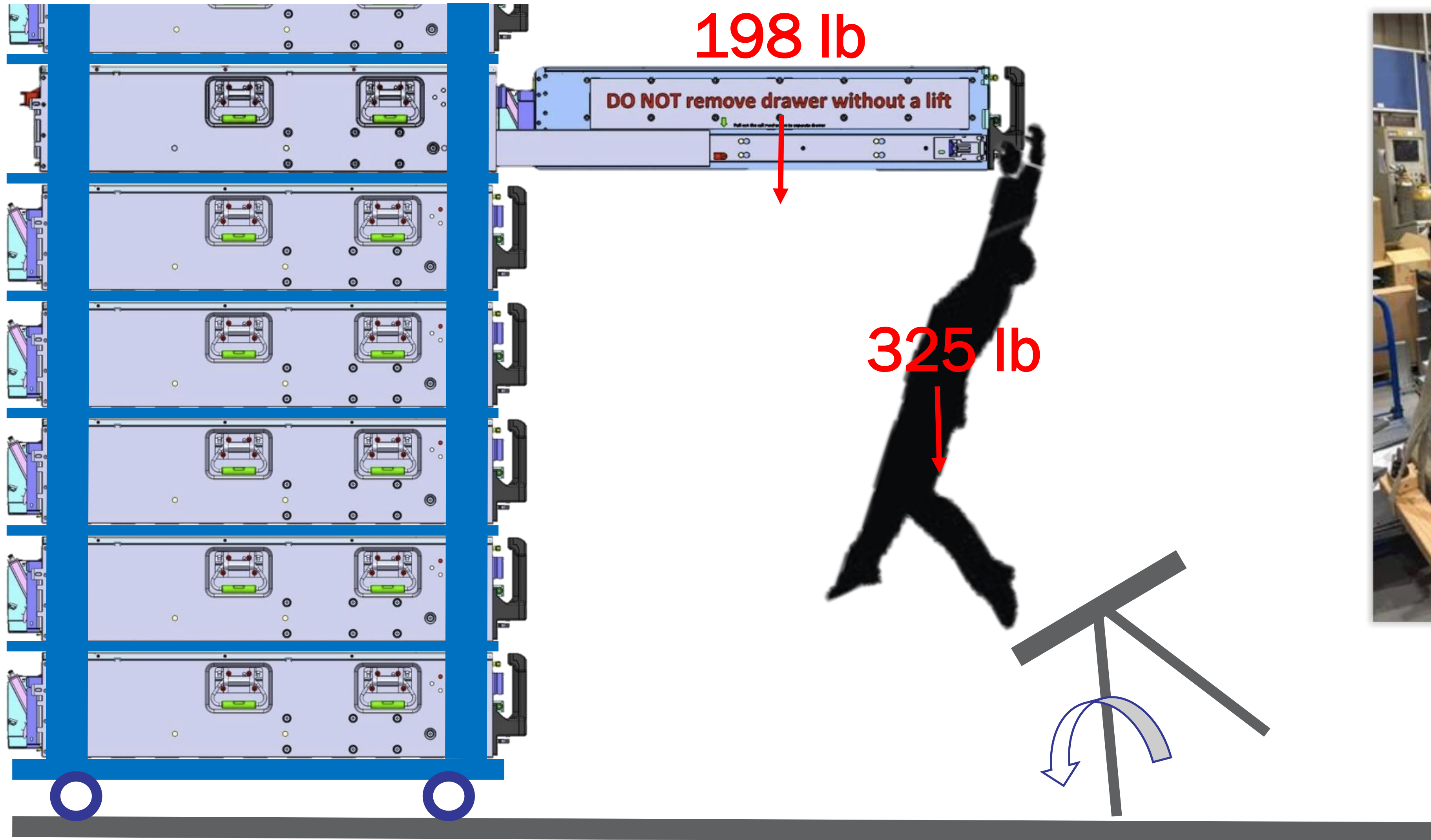


# System Improvements

Mechanical

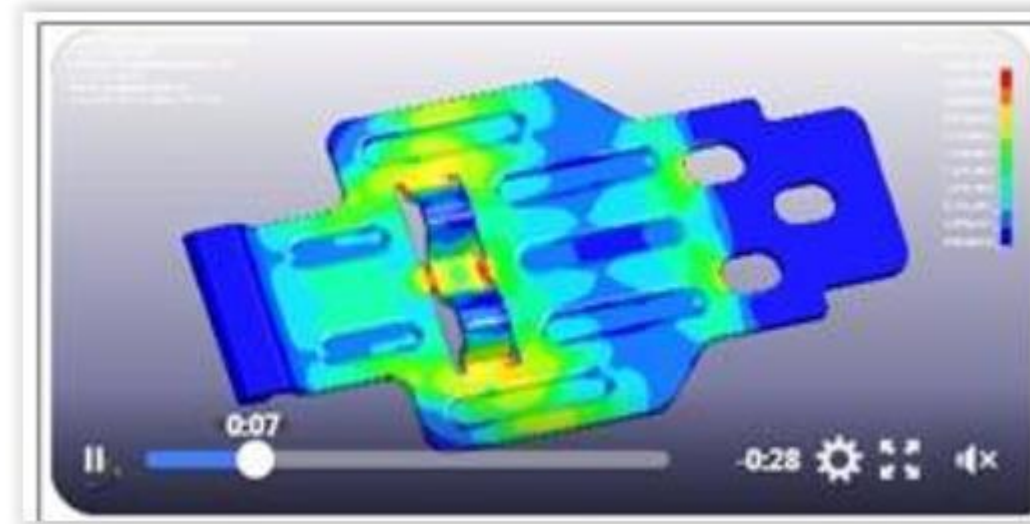


# Safety- robust rail





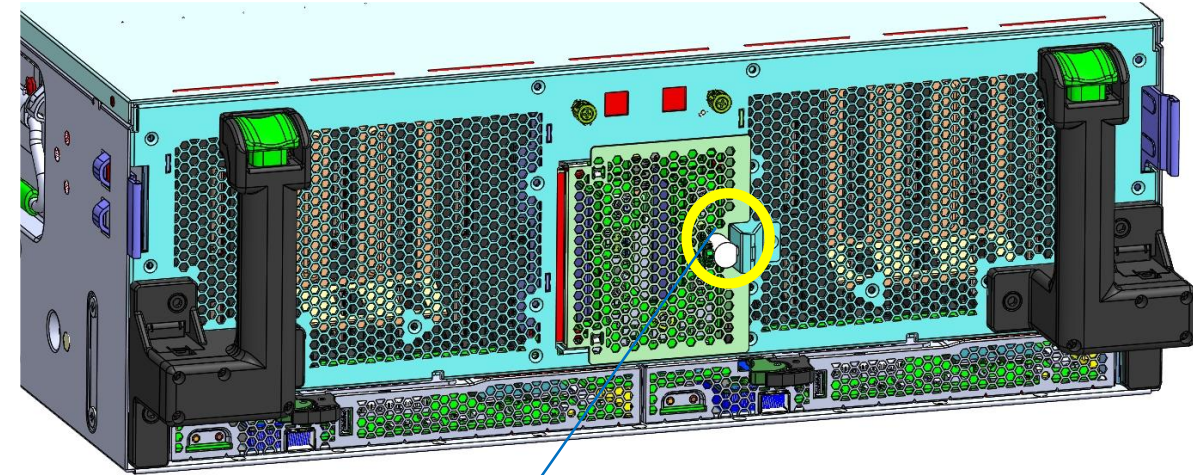
# Robust chassis rack lock latch





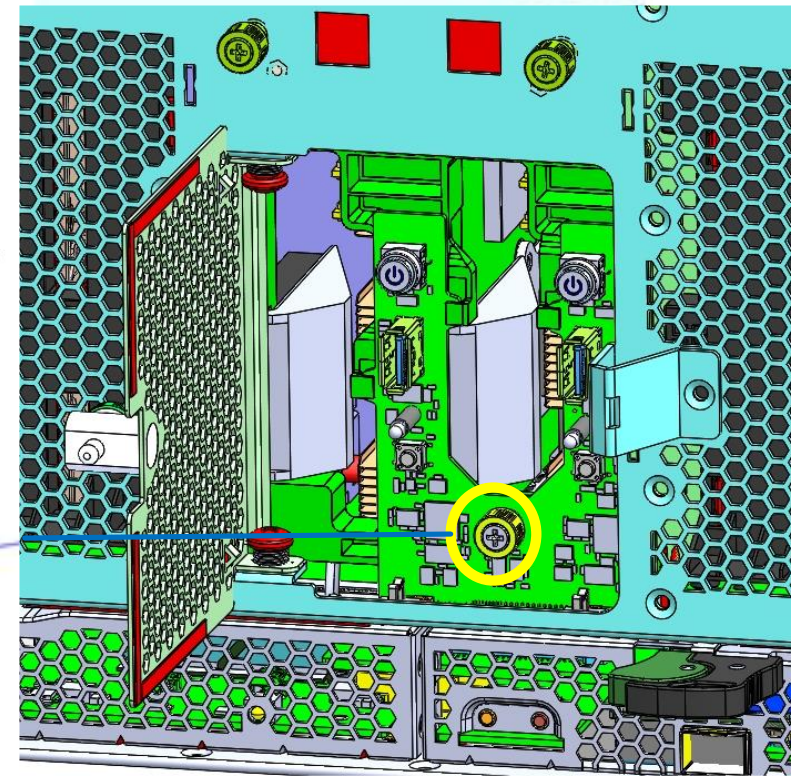
# Serviceability updates

A



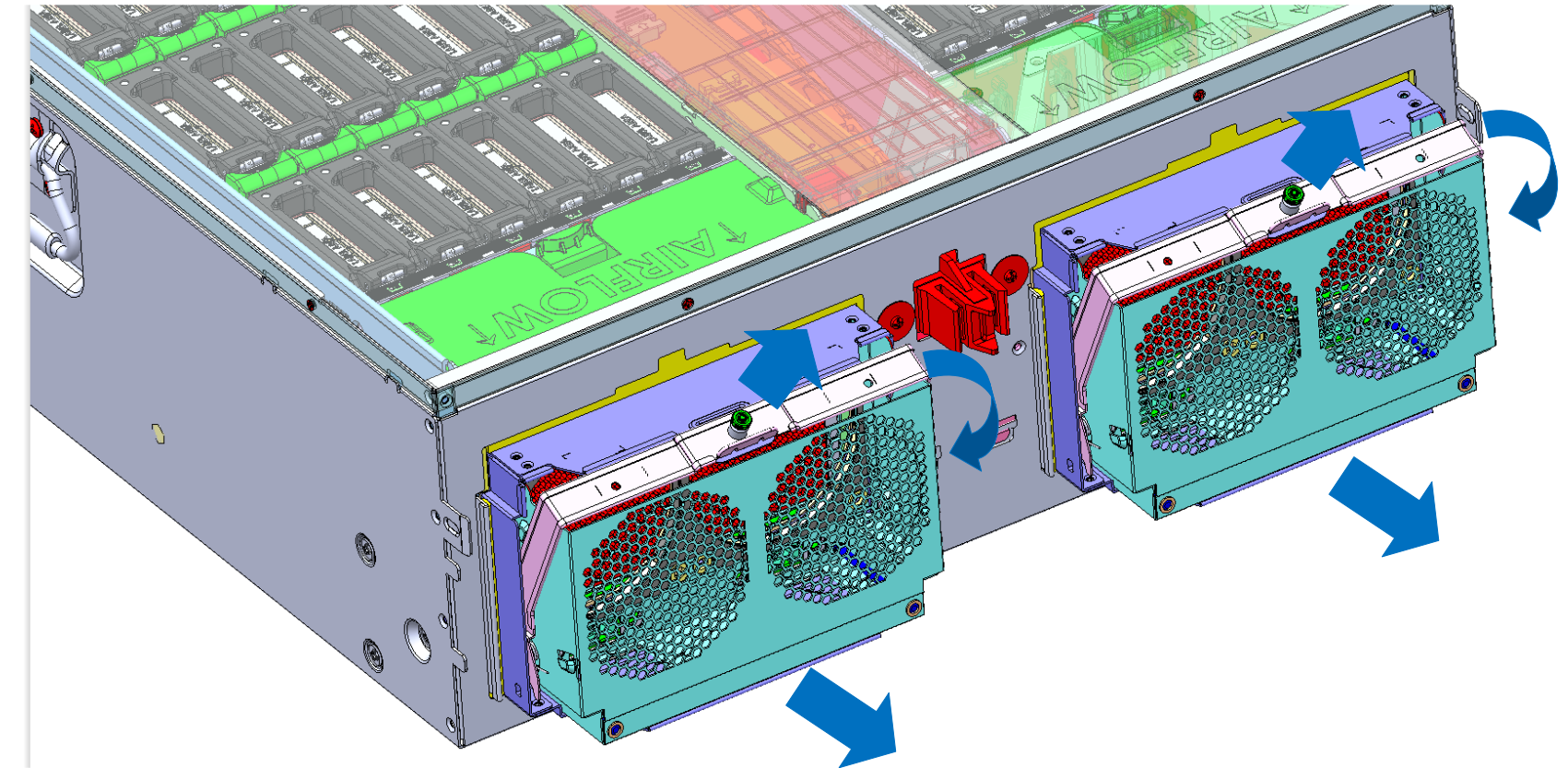
1. Pull the plunger on the door to the left to open the door and access the front panel board.

2. Loosen the thumb screw on the front panel board to replace/service it.



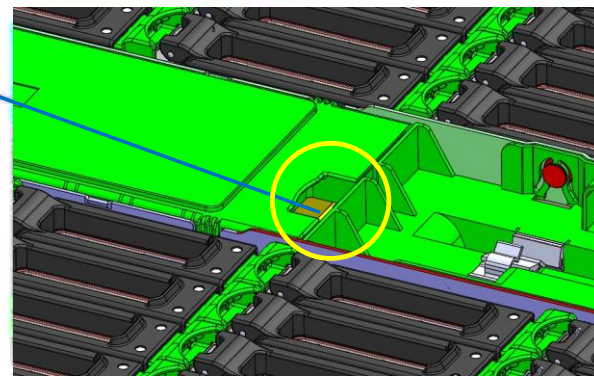
C

1. Pull the screw outward.  
 2. The fan cage handle will unlatch. Pull the fan cage handle downward to partly eject the fan.  
 3. Pull the fan cage outward to remove from the system.

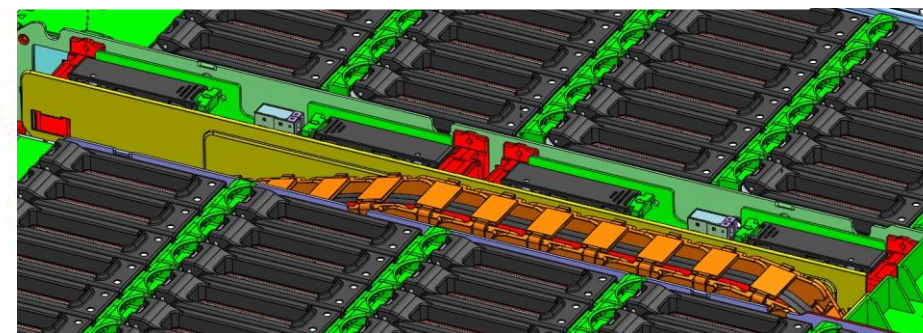
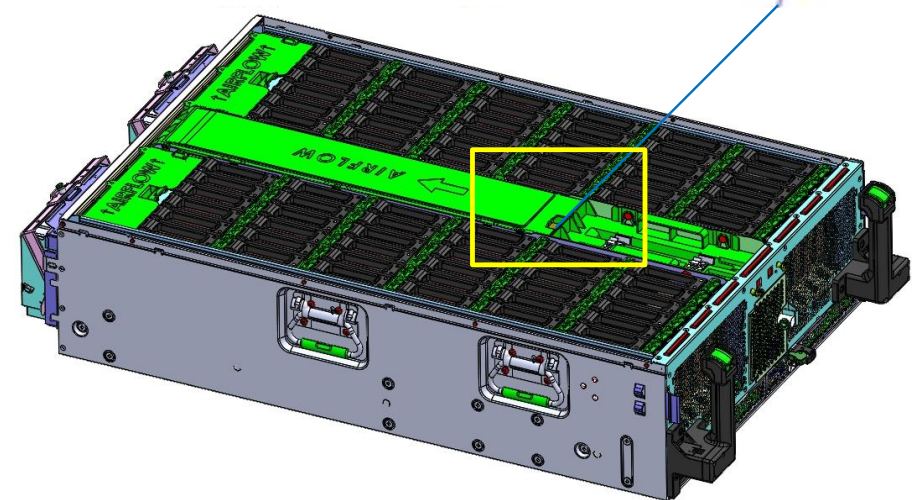


B

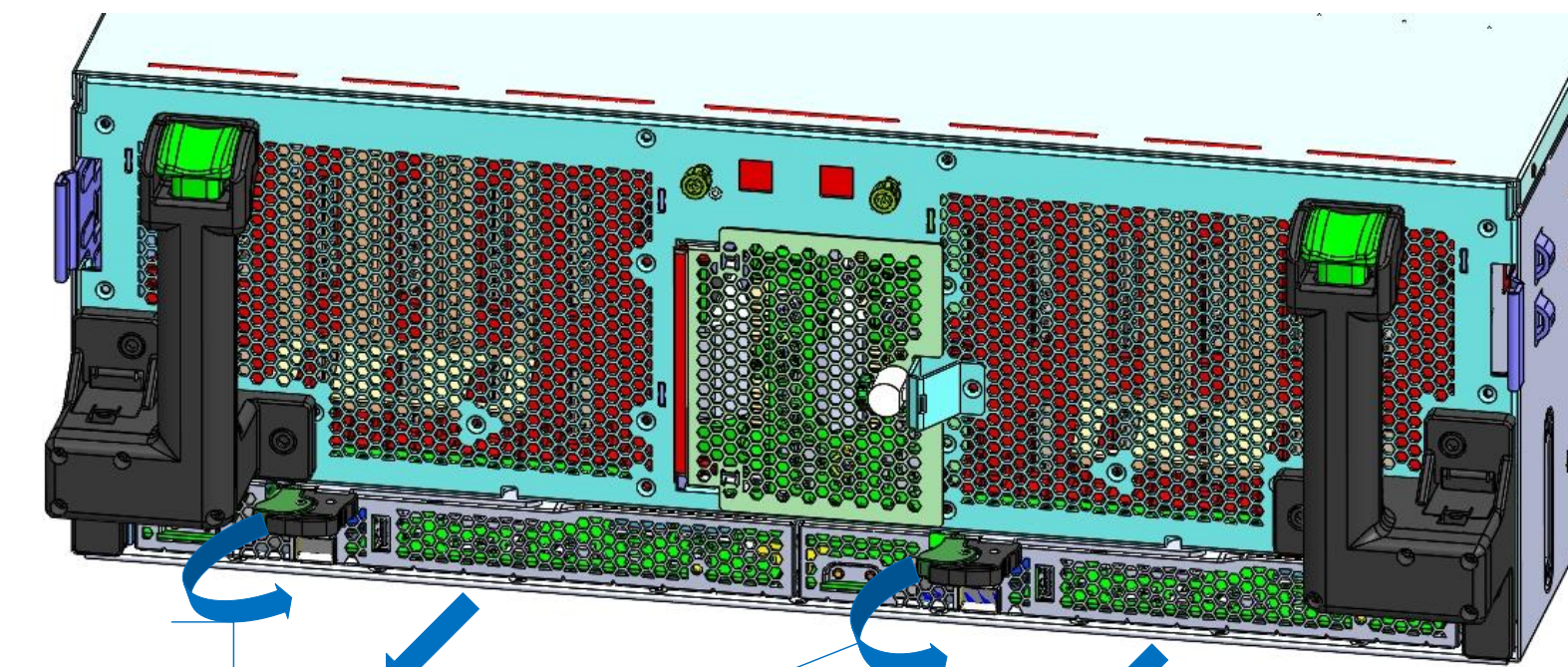
1. Gently pull out the plastic cover over the SCC cards slot.



2. Release SCC latch and pull SCC up.



D



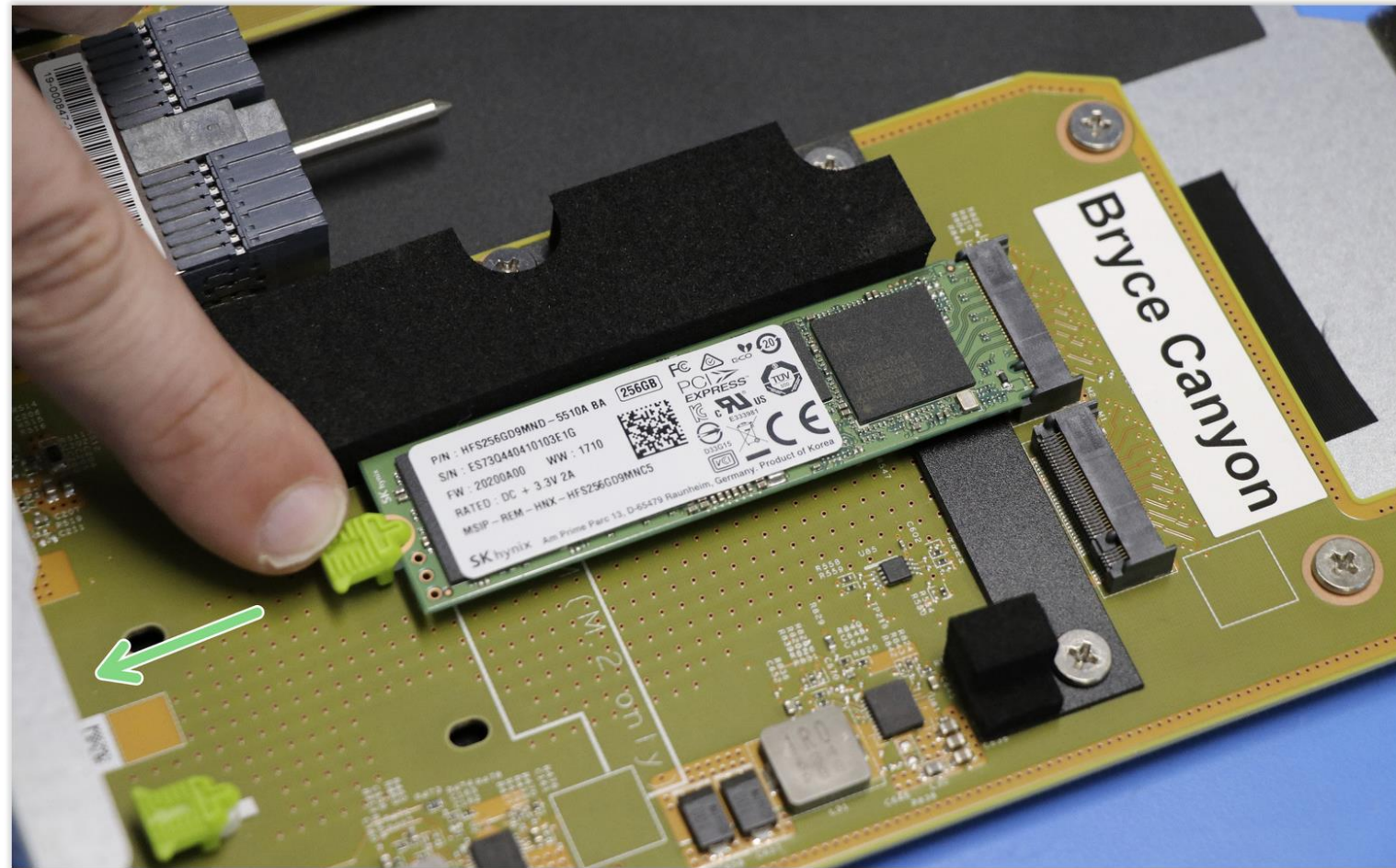
Pinch the latch handle and rotate the latch to release the IOM. Then pull the IOM out.

One hand tool-less operation

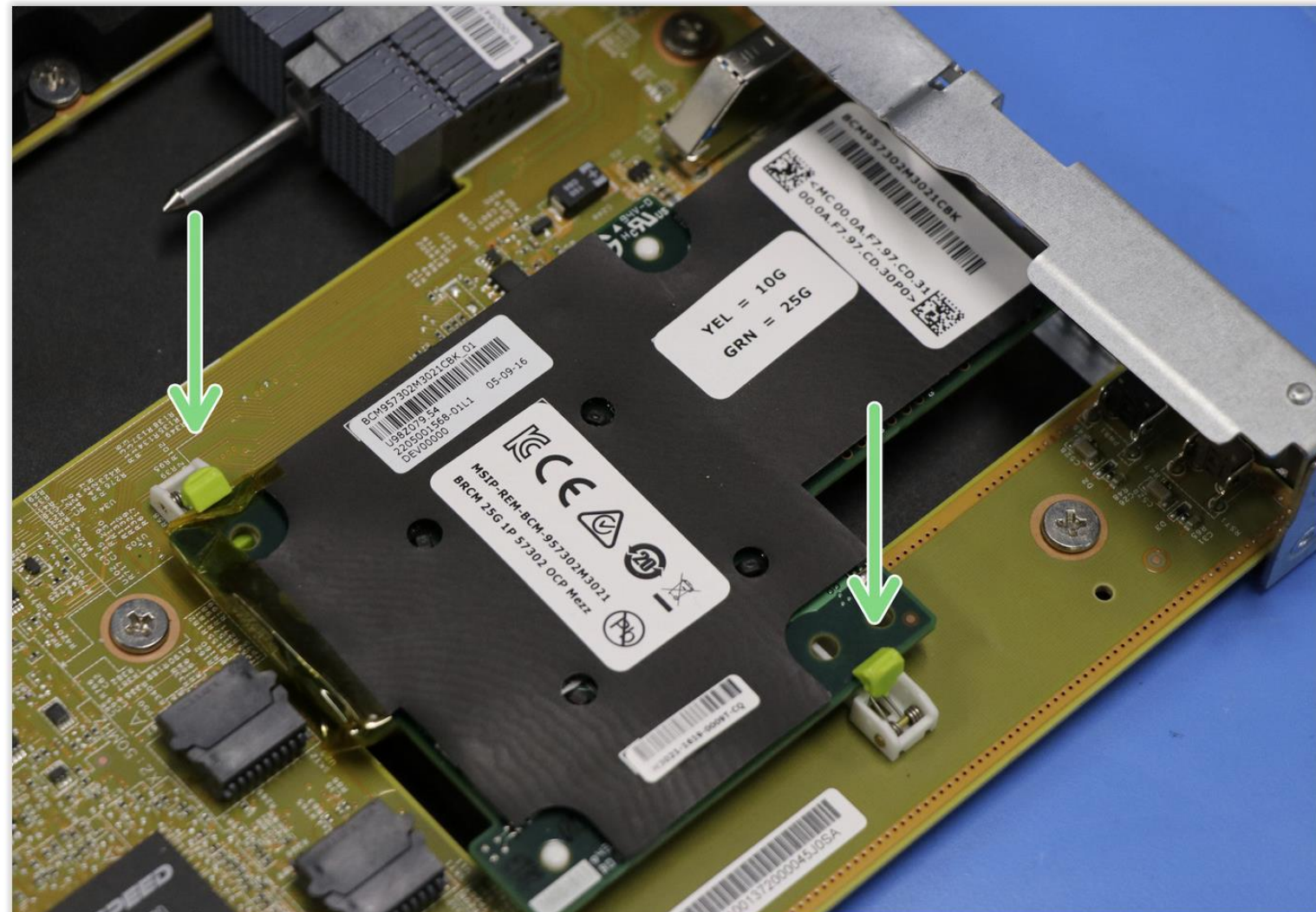


# Serviceability updates

A M.2 Card retention latch



B NIC Card retention latches



C Carrier-less HDD latch



D SCC Card latches





# Serviceability updates

Ergonomic thumbscrews



One piece backplane replacement





# Bryce Canyon OCP contributions

- Bryce Canyon system specification update
  - <http://files.opencompute.org/oc/public.php?service=files&t=ff9615864f33731e50c21a33f2656dc8>
- Electrical and mechanical design package release from our design partners
  - <http://files.opencompute.org/oc/public.php?service=files&t=d0cc07497b6f85c337a73eaa6225d163>
- OpenBMC Github release
  - <https://github.com/facebook/openbmc>



# OCP SUMMIT