OCP Engineering Workshop 25 September 2017| Dallas, TX

OCP NIC 3.0
Overview and Status

Jia Ning
Hardware Engineering Manager, Facebook

Agenda
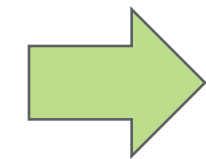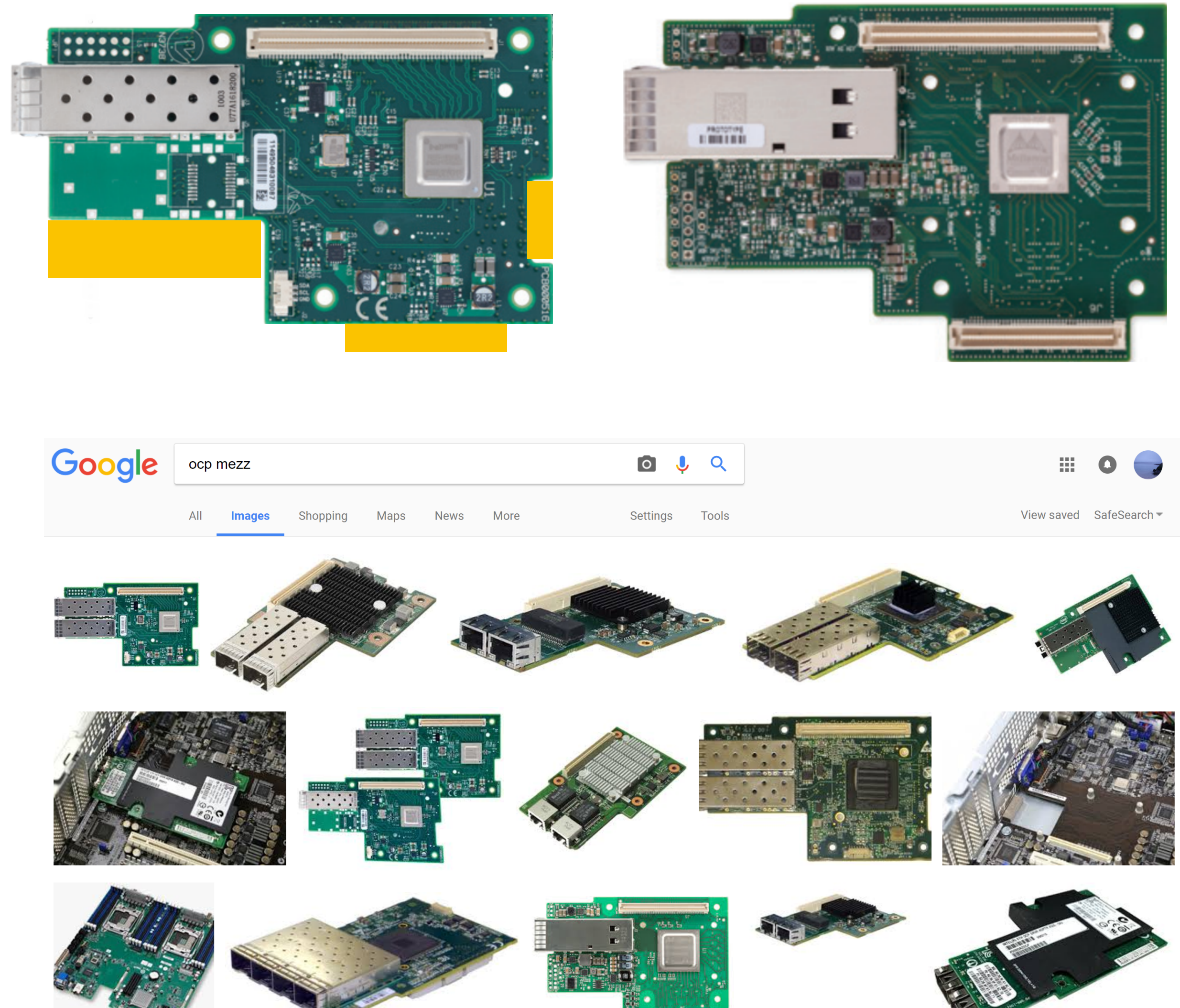
1:30-1:50	20 min	Welcome and General update	Bill Carter - CTO, OCP Archna Haylock - Community Director, OCP
1:50-2:30	40 min	Project Olympus update	Mark Shaw, Microsoft
2:30-2:40	10 min	Break / buffer	
2:40-2:50	10 min	Overview of OCP NIC 3.0 status	Jia Ning, Facebook
2:50-3:00	10 min	OCP NIC 3.0 - Thermal considerations	Yueming Li, Facebook
3:00-3:10	10 min	OCP NIC 3.0 - Mechanical update	Joshua Held, Facebook
3:10-3:20	10 min	Mechanical Mock-up review and demo	Joshua Held, Facebook
3:20 – 3:30	10 min	Break / buffer	
3:30 – 3:45	15 min	OCP NIC 3.0 Connector Update – TE	Arash Behziz / David Herring, TE
3:45-4:00	15 min	OCP NIC 3.0 Connector Update - Amphenol	Albert Chen, Amphenol
4:00 – 4:20	20 min	OCP NIC 3.0 - Dell – EMC update	Jon Lewis, Dell-EMC
4:20 – 4:30	10 min	Break / buffer	
4:30 – 5:00	30min	Pinout discussion	Paul Kappler, Intel & Jia Ning, Facebook
5-5:30	30min	NIC Management	Hemal Shah, Broadcom

Background

- OCP Mezzanine NIC v2.0

  - An **open** compact NIC module specification solves system configuration problem

  - Relatively **close** and **incremental** development process

  - **Healthy** adoption on "white box" market

  - Hits use case **limitations** – thermal, PCB placement space, etc.

  - NIC suppliers strongly asked for specification refresh in 2016/early 2017


Google   ocp mezz

All   Images   Shopping   Maps   News   More   Settings   Tools   View saved   SafeSearch

* Images from internet

Status

- OCP NIC v3.0
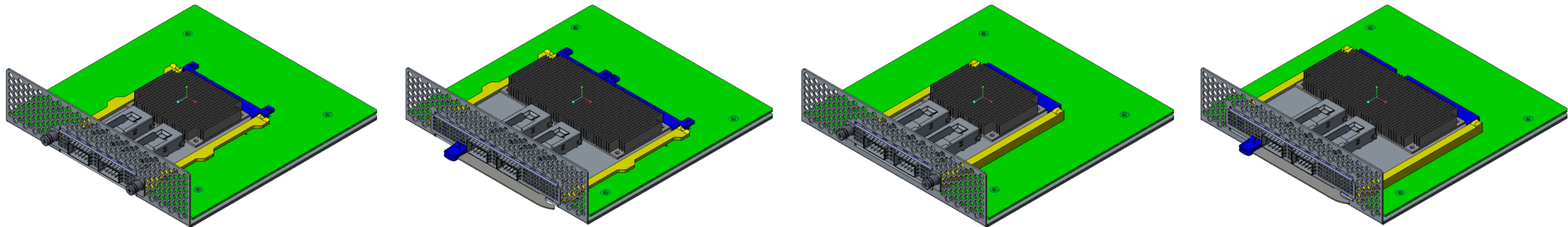
  - Community development since Feb 2017

  - Joint effort across system/NIC/Connector suppliers, and end users

  - Major challenges to solve:

    - Have backward compatibility
    - Manage NIC / Platform transition
    - Optimize PCIe Gen4 SI and routing / Provision for PCIe Gen5
    - Enlarge PCB space for Smart NIC use case
    - Allow higher TDP ASICs
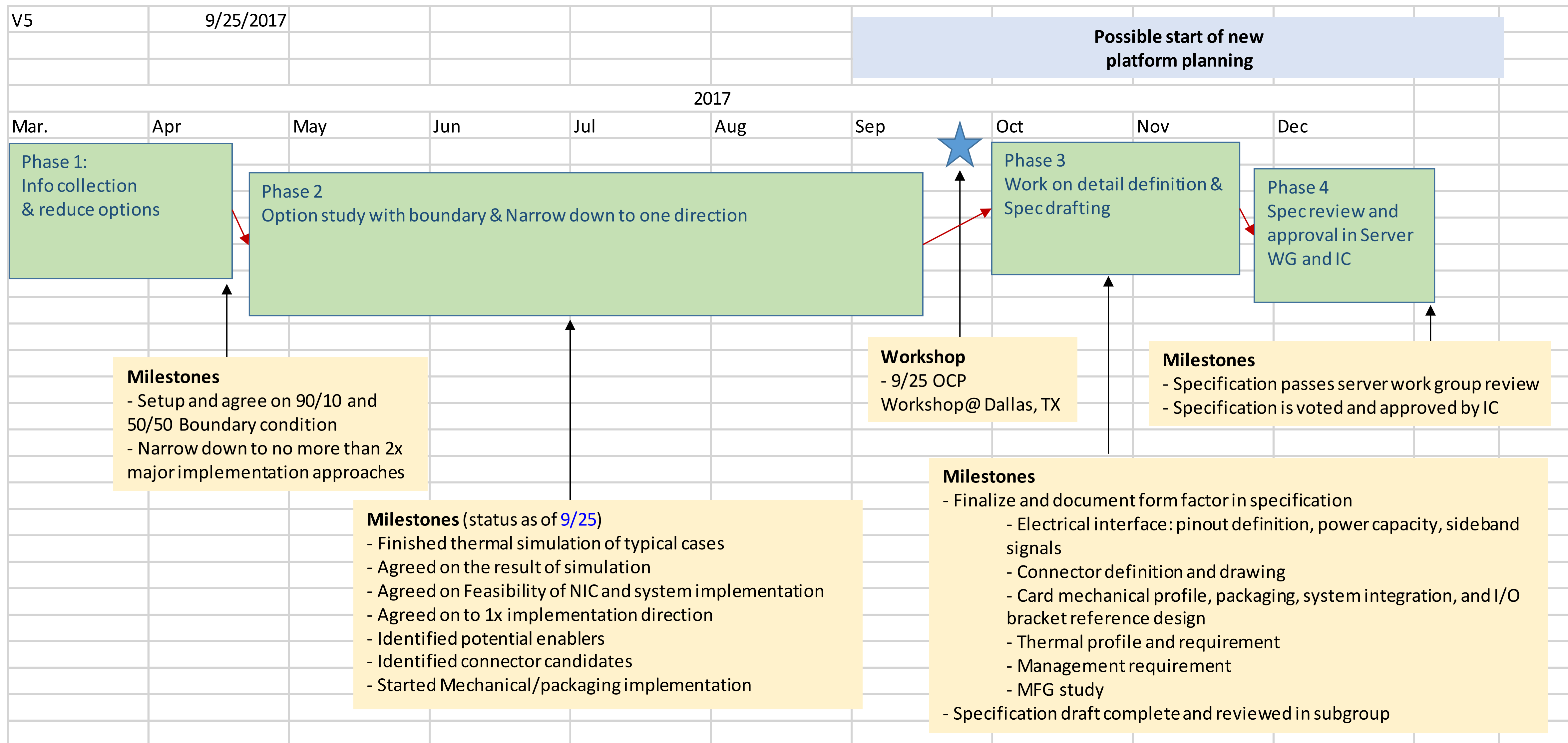   - Improve Mechanical interface

Status

Description of change made to Mezz 2.0	Status	Feedback	2.0 Compatibility
1 Add 16mm stacking height option	Close	NIC placement challenge has no improvement	Yes
2 Extend width	Close	NIC placement challenge has no improvement	Yes
3 Extend length	Close	NIC Placement challenge is partially addressed NIC PCIe routing challenge still exist	Yes
4 Move connector B to right edge	Close	NIC PCIe routing challenge gets worse	Partial
5 Flip ASIC to top	Close	Force tradeoff between system configuration flexibility and thermal	Yes
6 Move connector A to the left edge and keep same Y-location as Mezz 2.0	Close	NIC PCIe layout has crossing Possible backward compatibility for x8	Partial
7 Move connector A to the left edge with smooth PCIe lane sequence	Close	Lack of backward compatibility Most challenging for migration Being used in Simulation in June/July	No
8 Based on 7 and turn connector B by 180º	Close	NIC PCIe layout has crossing Possible backward compatibility for x16 with dual layout	Partial
9 2x Mezz 2.0 footprint side by side	Close	Limited thermal improvement due to connector placement	Yes
10 RA PCIe connector at side	Close	Stacking prevents 1U system to fit 1x Mezz + 1 PCIe CEM	No
11 Co-planner/ straddle Mount connector at rear	Close	Major pros (thermal and servicing) and cons(large cutout, needs extension for small card) identified	No
12 Right Angle Connector at rear	Close	Thermal concern of RA connector dams the air flow Stacking prevents 1U system to fit 1x Mezz + 1 PCIe CEM	No
13 Co-Planner / Straddle Mount Connector at side	Close	New option; Major pros (thermal) and cons(large cutout) identified	No
14 Straddle Mount / RA at rear; Expend towards the side	Open	- Confirmed on 9/20 as only option for OCP NIC 3.0 - Got most support from NIC / System / Connector suppliers - More details to be worked out	No
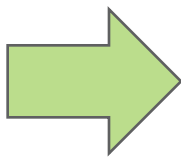
- Converged to Option #14
  - Does not have backward compatibility
  - 2x sizes of NIC modules
    - 115mm L x 74mm W1 for simple NIC
    - 115mm L x 118mm W2 for smart NIC
  - 2x connector styles on system
  - Straddle Mount / Right Angle
  - Tradeoff between Z-height and baseboard cut
  - NIC is compatible with both connector styles on system

Program Schedule

Possible start of new platform planning

2017

Mar.	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec

Phase 1:
Info collection & reduce options

Phase 2
Option study with boundary & Narrow down to one direction

Phase 3
Work on detail definition & Spec drafting

Phase 4
Spec review and approval in Server WG and IC

Milestones
- Setup and agree on 90/10 and 50/50 Boundary condition
- Narrow down to no more than 2x major implementation approaches

Milestones (status as of 9/25)
- Finished thermal simulation of typical cases
- Agreed on the result of simulation
- Agreed on Feasibility of NIC and system implementation
- Agreed on to 1x implementation direction
- Identified potential enablers
- Identified connector candidates
- Started Mechanical/packaging implementation

Workshop
- 9/25 OCP Workshop @ Dallas, TX

Milestones
- Finalize and document form factor in specification
    - Electrical interface: pinout definition, power capacity, sideband signals
    - Connector definition and drawing
    - Card mechanical profile, packaging, system integration, and I/O bracket reference design
    - Thermal profile and requirement
    - Management requirement
    - MFG study
- Specification draft complete and reviewed in subgroup

Milestones
- Specification passes server work group review
- Specification is voted and approved by IC

Near future focuses

- Discuss and finalize pinout

- Iterate and finalize mechanical and packing detail

- Perform more detailed thermal simulation, including optical module use cases

- Support connector enabling *

- Specification drafting

- Specification voting in workgroup, and approval in OCP IC

- Sub-Workgroup intends to collect and converge requirements and work with connector suppliers for enabling on good will.
- Sub-Workgroup **DOES NOT** guarantee the enabling of any specific connector in specific timeline.
- Sub-Workgroup intends to document P/N and availability of specific connector only if it is public information.
- Communication of specific connector P/N and availability information <u>prior to public release</u> will be from connector suppliers to Sub-workgroup or individually companies.

Resources

- OCP Server workgroup – NIC subgroup Wiki

  http://www.opencompute.org/wiki/Server/Mezz

  - Regular update of status, trend / Archive of discussion, document and presentations from community

- Bi-monthly call

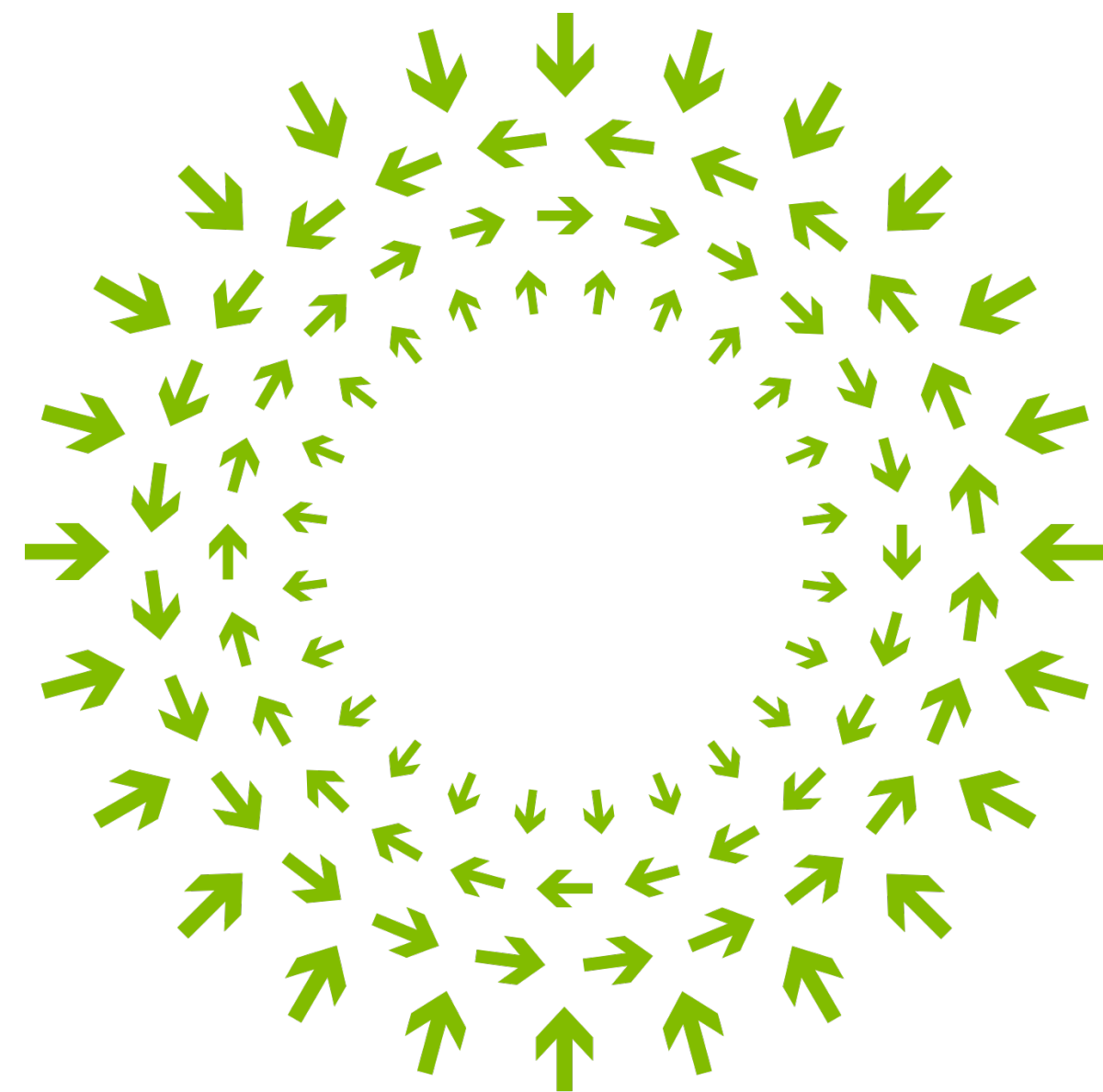 opencompute.org/community/ocp-calendars

  - 1st and 3rd Wed 9am PDT / Key updates to community and feedback gathering / Presentation from stakeholders

- Mailing list

  http://lists.opencompute.org/mailman/listinfo/opencompute-mezz-card

  - Meeting notice and agenda / offline discussion

- Specification drafting review meeting (coming)

OPEN
Compute Project