OCP Server Working Group meeting, 7/15/2015

Attendees:

Mark Shaw, Microsoft

John Stuewe, Dell

Bob Ogrey, AMD

Jia Ning, Facebook

Marc Licciardi

Martin Goldstein, Microsoft

"RR" Ron Renwick

"John"

Fred Worley, Samsung

Erwin C, Emulex

Bill Carter, Intel

Dan, JP Morgan

Paul Hartke, Xilinx

Minutes by Fred Worley

Agenda:

ACTION ITEMS:

1. AI (Server Working Group): Provide feedback on Yosemite multi-node server specification posted 6/17/15
    a. http://files.opencompute.org/oc/public.php?service=files&t=cfc45241f28367e4eff7fb60d6f69f3f
2. AI (Server Working Group members): Please provide appropriate review of the new thermal addendum
    a. http://files.opencompute.org/oc/public.php?service=files&t=7d52c3023a73f5c4b7b37aee2b80879a
3. AI (Jia Ning/Facebook): Change title of the thermal addendum document to clarify that the document is an Addendum
4. AI (Jia Ning/Facebook): Add the legal requirements, etc., from the standard document template to the addendum add-on-card thermal interface spec
5. AI (Jia Ning/Facebook): Help improve the Yosemite spec, and specs in general, to make the interface requirements unambiguous (or for any interface that Facebook defines)

Meeting Summary:

- Facebook Add-on-Card Thermal Interface Spec
  - o Change name to be clearer as an addendum to the v3.1. Take as a pair?
  - o Wants change from 3.0 to 3.1 reviewed. Implementation not in 3.0 or 3.1 appendix. 3.1 directs to the appendix. Want to retire the 3.0. Jia to give a rev to include in. Module can be used on other boards. Concerned that this card is going to rev on a different cadence and gives guidance to vendors to build different cards…
  - o How do you test this system? Expect vendors to implement and test this part.
- Micro-server/Honey Badger/Panther+ I/F
  - o Feedback to Mike Yang?
  - o Management i/f was lacking, unable to build to.
  - o Management is on the baseboard, but no specification
  - o Bob to review the updated management spec. Review next time in the committee.
- Yosemite Multi-node server specification released on June 17th, 2015
  - o Comments on the need for detail in the management interface of the specification
    - ▪ The specification focuses on the SOC board, but the management fuctions are provided by the baseboard
    - ▪ The baseboard doesn't have a specification, so this leaves too much about the management of the SOC boards unspecified
    - ▪ Request for future versions of the specification to address this point and add more detail on management of the full solution
    - ▪ Request by Facebook that feedback and suggestions for improvement of the specification be provided to Mike Yan at Facebook
  - o Additional comment that there is some question on how the cards communicate with each other
  - o Request for further review of that specification by the Server WG
- Active project review
  - o Everything posted to the web page has been accepted by the IC accept:
    - ▪ There is an addendum to the thermal interface posted by Facebook
      - • There is this add-on thermal interface card that covers a gap in the implementation
    - ▪ Please review
      - • http://files.opencompute.org/oc/public.php?service=files&t=7d52c3023 a73f5c4b7b37aee2b80879a
    - ▪ Request that the title of the document change to reflect that it is an addendum
  - o Peripherals:
    - ▪ All of the posted peripheral specs have been accepted by the IC

- o Other EIA-310D compatible specifications
  - All posted have bene approved
- o Things heading to IC
  - Open CloudServer power supply and solid state drive
  - Panther+ contribution has been submitted by Quanta – including electrical specs for Avaton microserver
- o Things requiring work
  - 1 socket multi-node platform, Yosemite platform that goes with it
    - Note discussion on this topic, below
- Discussion of 1-socket multi-node platform issues
  - o For AMD Seattle, had USB and PCIe clock pinouts that conflicted with other microserver cards
    - Alignment needed with an accepted spec, like Panther Plus
    - Noted that even with an updated specification for the AMD Seattle spec, adding a USB controller to the design, the system management interface will still not be compatible with any of the storage boxes
    - This is a fundamental issue with the lack of detail on that first generation storage spec
    - The forward implementation should be consistent with the newer version of the spec
    - Bob wants to update to the new 1S management I/F.
    - Panther+ does not have bridge IC.
    - FB does not intend Panther+ to work in Yosemite system with depopulated connector.
    - It appears that Storage Vault is the only consumer of the micro-server and AMD Seattle isn't going into SV. Bob's thinking that AMD Seattle spec to be directed to the new 1S spec.
    - Should there be commonality?
    - Moving forward, would be good to make it so that future commonality is possible.
    - 1S spec calls out the commands and protocols for IPMI over i2c
    -
- Discussion of how to have a common single-socket board design for server and compute applications
  - o Moving forward, plan is with the new spec to have the same types of cards go into different systems
  - o However, requirements for going into the compute node is totally different from the storage spec
    - The new design with instances of x16 and dual x16 will be different
  - o Prefer to be compliant with a new spec, rather than patch the old spec
  - o Management interface differences also need to be resolved

- The edge connector for the microserver and one socket cards are electrically compatible
- But how the nodes communicate and the management interface they use are not compatible
  - o Detail:
    - Jia Ning, Facebook: the compute node leveraged from earlier efforts, for 20-30W SOC
    - The one socket spec is for multi-node but supports 45-60W SOCs
    - They are for different applications
    - It happens that the edge card is similar, since we leveraged some work from the [earlier effort]
    - But the management interface itself is not the same
    - The 1-socket card has the bridge IC on the card, while the [smaller board] card does not
    - Facebook does not intend to change the Panther cards to be forward compatible to the one-socket Yosemite
    - These are considered 2 different products, 2 different applications
    - Clarified that if you depopulate the rear connector on a Yosemite, panther plus still won't work in it
- Discussion of how to have a common CPU board definition for compute and storage
  - o Noted that there is no requirement for a common compute and storage platform from Facebook
  - o Comment that there is interest (generally, not from Facebook) for using 1-socket in both uServer and storage apps
    - If these concepts are taking forward, they likely won't be in the uServer format
    - Suggestion to use Yosemite multi-node platform for this class of platform
    - More applications for single-socket boards (e.g. in server and storage platforms) will increase volume and adoption
  - o Question of whether there are platforms outside of OpenVault, including OpenServer, where processors like the AMD Seattle would be deployed
    - If not, could standardize on Yosemite as the platform moving forward
  - o Question of how to leverage work on single socket for deployment in Honey Badger storage platform
    - Noted that Facebook does not require alignment between these platforms
    - Don't want to overly burden the implementations
    - Can take advantage of leverage between platforms, but don't want to force one to include unnecessary components for that use case to meet the needs of the other
    - Noted that there is no reason a company couldn't propose a version of Honey Badger that would accept a common board with Yosemite
    - Such a board would need a full management interface definition
  - o How to have a common management interface between compute and storage

- Discussion on interface design as a WG activity
  - If we do a really good job defining interfaces, we then allow the industry, suppliers, end users to define implementation specific or application specific nodes that plug into that interface
    - We get more choices as a consumer
    - More important to work on those interfaces than in generating product collateral and the designs
    - If the interface is really well defined, you can have 3-4 or even 10 different add in cards that plug into it
    - Achieves the goal of having lots of choices
  - Interfaces have 3 parts:
    - Mechanical
    - Electrical
    - Protocol
    - All three have to be defined
  - What interfaces have been defined to date:
    - A socket specification is being developed for Yosemite
      - Jia Ning of Facebook offered to help improve the spec to make it unambiguous (or for any interface that Facebook defines)
    - The one socket spec is still a work in progress
      - Need definition of the protocol, what is happening over the I2C
      - The IPMI command is specified with the command format for how the baseboard communicates with the command interface
      - Facebook requests feedback from the community to make the spec more complete
      - Comment that the spec is a lot better in its current version
      - Spec is posted under "Specifications requiring Additional Work" – the one discussed today is v0.3 of the OCP 1S Server Design Specification (posted 2/18/15 by Facebook)
        - http://files.opencompute.org/oc/public.php?service=files&t=433afe70a17e23746c070db483034d58
    - All the OCS projects defined interfaces to the blades and management
- Discussion of document versioning standard
  - Significant discussion of how to organize and name the documents and specifications within the OCP Server project to clarify how the documents are correlated and to standardize the meaning of version numbers
  - While this topics was discussed at length, no formal decisions were made. Proposals are listed below:
  - Suggestions that:
    - When a document is accepted by the IC the version number of the specification be changed to v1.0 (define v1.0 to mean acceptance by the IC)

- Discussed methods to handle the need for an addendum, and merge an amended version of a spec back into a later version of the base spec
  - Example: the thermal spec addendum for the OpenRack Server v3 specification specifies the thermal requirements for add-in cards, not for the server itself
  - It is not clear, therefore, how to roll the requirement back into the base spec
    - Facebook prefers to leave the add-in card thermal requirements as an appendix to the 3.1 spec, not part of the base 3.1 spec,
    - Having the document as a separate appendix allows the thermal requirements for add-in cards to be referenced by other server platforms
    - Desire to allow the card vendors to only track the thermal requirements spec, not to have to track unrelated changes to the base spec, or to track multiple specs if the requirements are common across server platforms
  - However, it would be preferable to have a v3.1 of the Open Rack Server spec that incorporates all previous specs and addendums that would supersede any previous spec, allowing the 3.0 spec to be retired
- Discussion of whether the specifications should be modular (e.g. a thermal specification that can be referenced by multiple specs) or stand-alone
  - Issue discussed at length but not resolved
  - Request made that if documents are intended to be modular, the naming and context of the documents needs to be significantly more clear
  - The issue of how a particular server, card, etc., would be certified to be compliant with a modular specification was raised – not clear how to specify what set of specifications would be needed to be compliant
  - Noted that the testing issue is broadly unresolved as well – when a specification is submitted, it does not typically come with a test document that would certify compliance
  - Request for greater clarity in the testing process
    - This is likely a significant undertaking, but one where the Server WG could provide some useful guidance
- Vendor NDA Submissions
  - One vendor has a specification for review but still requires an NDA to review it
  - Mark Shaw, Server WG co-chair, will ask for clarification of their intentions
  - The need of companies to protect their product plans while building products for an open architecture is an example of the importance of having complete interface specifications
- Future meeting plan:

- o Focus on 1-socket server and multi-node platform in next meetings, review specs in greater detail