In general I agree with the Katja, Mark, and Lea the paper is well written, and overall the results are presented in a clear manner; however the paper is pretty short and overall I think there needs to be more background information provided to help non-experts understand the material presented. Especially since I think (maybe?) that this might be the first publication with results from a test beam of a PS prototype, so it would be good if a non-expert could already get a good overview of the module without referring to the TDR/publications from the ASICs or hybrid design teams.

Section 1 :

Perhaps the standard tk layout of the tracker indicating the location of the PS and 2S modules would be good to include. And possibly also one of the rendered images of the PS module design?

General comment throughout the text, whenever possible please try to be consistent in the use of hit/cluster/centroid to describe the L1 information coming out of the SSAs.

L4 : Please be a bit more precise with regards to how this compares to the current LHC (xx fluence, xx current occupancy)

L11 : Please define all acronyms (MPA)

L13 : Its not clear to me from the text what combined means here. I know its into stubs to be forwarded to the CIC.. but a casual reader will not necessarily. Also, if you plan on keeping the reference to stub and L1-trigger path in the rest of the document that L1 trigger and L1 accept here is confusing.

Section 2 :

Please consider taking some time to describe the data flow on the PS module in a bit more detail, preferably with the help of a diagram illustrating the two data paths. It will make the sub-sections describing the SSA module much easier to follow/understand.

In addition, I think it would be useful if you could expand on the triggering and clocking scheme used in the test beam in a bit more detail. Since you do perform a timing scan and refer to latencies/timing later in the document a small diagram in this section could help make the remainder of the paper easier to follow.

L27 : If you know what the value (in electrons or fC) is expected to be for a fully depleted PS strip sensor then I would suggest including it here.

L29 : This is the first time you refer to a ‘stub’ in the text .. anyone outside of CMS OT upgrade will not know what this means.

L37 : The SSAs do not transfer stub information to the MPA, but simply the centroid data. The logic required for forming stubs with the strip and pixel clusters is in the MPA rather than the SSA. I am sure that you are already well aware of this but I personally find way this sentence is written to be a bit confusing.

L48 : downstream?

L49 : I find this a little bit terse.

Figure 4: Personal preference, but in the caption I would rather that ‘high precision’ be quantified. I know its in the text but … sometimes you scan a paper quickly and having the information in more than once place does not hurt.

Section 3 :

I would find it very helpful if you could add to this section an illustration which shows two strips , the different mechanisms by which charge is collected by the strips, and a few equations that

describe how the charge collected by a given strip is expressed in the model. I know its all in the text, but I personally find it easier to follow the text when there is more information provided. Also, it would be nice if you could show an example output of the simulation for what you expect the nominal parameters to be based on the detector that you are studying. This already prepares the reader for the type of plots you collect at the beam test and I thin can really help make the discussion in the paper much easier to follow.

L52 : Which data?
L55 : Please define U explicitly " applied voltage U, and depletion voltage Ud "
L66 : How is the noise of the sensor determined?
L66 : Its not clear to me, perhaps simply because I need to spend more time trying to draw everything out, if its necessary to use the data collected at various angles and thresholds to constrain the model? Could one dataset suffice?

Section 4 :

L74 : Please try and be consistent about whether the signal and threshold are described in fC or electrons. I know I can easily convert betweent the two but its much more enjoyable to read the text and focus on the work done when the units are the same throughout.
L76 : If you have a plot showing the distribution of cluster size then please include it here.
L78 : I think its clearer if you just use 'alignment' here rather than calibration.
L82 : Please explain here how you will be defining the efficiency in the analysis.
L83 : If possible, please include an illustration of this with all the angles/coordinates clearly labelled.
L91 : Please include either a diagram or an equation that helps the reader follow what you are describing in the text.
L95 : If possible, then perhaps this can be illustrated with a plot that showing the value of the minimized parameter as a function of z?
L109 : See previous comment, this should be really already have been defined.
L111 : Do you have any measurements that support these assumptions? If so please include them here.
L113 : How many of these timing distributions are well known? And how many are not well understood? Where do they come from?
L117 : I think if you go to the effort of describing the shape of the data shown in Figure.4 then you should also explain why no attempt was made to fit the shape. I would also explain that since this is binary readout in fact what you are doing is then taking the convolution of the ToA and the response of the front-end and applying a cut at a fixed amplitude.. which perfectly explains the shape you get.
L119 : An additional sentence explaining why it is important to maximize this would be good.
L124 : Do you mean that the alignment is repeated for each angle, or for each threshold? I would also suggest that you simply refer to it as alignment rather than 'alignment calibration' as the latter is a slightly unusual way of phrasing it.
L136 : Would it be possible to expand a little on how this can be done? Did you attempt to do this with the test beam data? If not .. then perhaps explain why not.
L140 : Perhaps this is obvious, but do you understand why the efficiency is lower for the second threshold?
L146 : If I refer to Figure.8 I can understand where this 14.7 um comes from. But it would be easier to follow if this was described in the text, then repeated in the figure.
Figure 8 : Why is there some structure in the distribution of 1 hit clusters at alpha = 8? its weirdly symmetric..
Figure 9 : Personally I think that labelling this axis with Efficiency is a bit confusing - because if you look at it quickly you assume that you are very bad at detecting two strip clusters away from the center, and at detecting one strip clusters right at 0. Which isn't the intention - since in

fact the black shows the overall efficiency and the red and the blue show the fraction of one and two strip clusters. Perhaps a better way can be found to label the axis/make the plot so that its a bit clearer.

Section 5 :

This is a bit.. terse as a conclusion. Perhaps include a table that shows the properties of the DUT studied with the data collected at the test beam and compare those with the specs? Also - given that you spend quite some time describing the simulation used to analyze the data collected a sentence or two about its use would also be useful.