Least likelihood combination experiment, explicit, investigate, pair, encounter a certain range, sufficient, collaboration, chat, evaluate, delivery, simply, integrate

least likelihood,combination,combination, experiment, explicit at least combination experiments likelihood explicit pai, investigate, investrage

align with, strategy, forecasting, regular,doable, stakeholder, for instance, slides, aligned with, assign appropriate engineer to fix, fincal

I think it would be better to align with you about HW test coverage strategies before further forecasting.

At least, combination, experiment, explicit, investigate, pair, encounter a certain range, sufficient, collaboration, chat evaluate, delivery, simply, integrate. Fincal,for instance, slides, aligned with, workout in total, automation test infrastructure, clarify,overlop clarify, estimate

Regular 规律的 doable 可行的 stakeholder 利益相关者

instances 例子

slides 幻灯片

aligned with 结盟

assign appropriate engineer to fix issue

Fiscal 会计的， 财政的

in total

automation test infrastructure

work out 实现

For instance例如

clarify 澄清

could you please help to clarify versions of "BSP"

is it the correct one

+ @DL-SWQA-SH folks for viz.

overlap 重叠的

estimate 估算

assume 假设

adjustment

stack

involvement参与

has an AI to coordinate with 配合

on a wider range of Windows systems

subset

The goal is to run

sanity 健全的

o The testing strategy

mentioned 提及

require several continuous hours

attack involvement, on a wide range of , subset, the goal is sanity, strategy, mentional, challenge, in previous, tobe leveraged, limitation, time frame, regular, capacities, presentation, optional behavior, demand, confirmation, forecast,functionality,

Challenge 挑战， 艰巨任务， 质疑

chan

in previous

to be leveraged

Whether it’s a known issue/limitation

in the following time frame?

regular

capacities 容量，能力 presentation 演示

At least 至少 Optional behavior行为 Demand 要求 confirmation 确认书 forecast 预测

Functionality 功能性的

Combination组合

Experiments实验

Likelihood可能性

Explicit 清楚的

Pair 配对

Investigate调查，研究 encounter a certain range sufficient, collaboration collaboration

Encounter 遭遇 collaboration determine, pick upcombination

a certain range在某一范围内

 Sufficient 足够的

Collaboration 协作

Determine 确定

Pick up 获得，捡起 aligned to leverage conclusions conclusions, evaluate evaluate, delivery delivery

aligned to 对齐

Conclusions 结论

Chat 聊天

Evaluate 评估

Delivered 递送，发表

Simply 简单的 significant simply delivery conclusion aligned to pick up determine collaboration collation sufficient a certain range encounter encounter, significant, catch up measure,scenarios scenarios, survey, survey, concern, concern. Trade off pipeline, staff,scenarios,

Investigated pair, catch up, measure,scenarios, survey, concern, trade-off. Pipeline, finalize escape, upcoming, unified,

Significant 相当数量的，重大的

catch up 赶上

Measure 测量

Scenarios 情节，脚本

Survey 调查

Concern 顾虑

Trade-off 权衡，协调

Pipeline 管道

Finalize

Staff 职员 pipeline trade-off concern survey scenarios scenarios measure catch up escape upcoming, unified, stack, leverage exclude the intention of assume,

Escapes 逃逸

Upcoming 即将来临的

Unified 一直的

Stacks 堆叠 (leveraging Ubuntu 2004) 杠杆 stacks leveraging leverage unified unified upcoming escape staff finalize the intention of exclude, handy,

the intention of 意图 the intention of

exclude排除exclude assume

assume 假设 assume bandwidth handy specialize bandwidth, specialize evaluate,

non-trivial 非平凡的

bandwidth 带宽 bandwidth

specialize 专门研究 specialize

evaluate评价

handy 手边的，有用的

handy canidiate onboard

variability 可变性

swap to 交换到

identify 识别

Raising priority to

a wide range

spare 空闲的不用的

Colossus 巨人

There simply is not enough "spare" bandwidth available

Hiring:

Manager candidate:

3rd round interview last week, pending feedback from interviewers.

Next week: none scheduled.

IC candidates:

Xuejiao Wang onboard date: 5/16.

5 phone (1st round) and 2 video (2nd round) interviews last week.

Next week: 2 video (2nd round) interview scheduled.

Interns: 1 offer in flight.

cudnn\_level\_tests\_L3:   significant testing blocked by [3636651](https://nam11.safelinks.protection.outlook.com/?url=https%3A%2F%2Fnvbugswb.nvidia.com%2FNVBugs5%2Fredir.aspx%3Furl%3D%2F3636651&data=05%7C01%7Cwilliamy%40nvidia.com%7C9d566335b54b40e4ede508da35d7bca2%7C43083d15727340c1b7db39efd9ccc17a%7C0%7C0%7C637881498070707644%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=QwrdStFhf3ON3rZsRo%2FaycMGCIXUZcFdQUg5tIsPqOI%3D&reserved=0) and [3612577](https://nam11.safelinks.protection.outlook.com/?url=https%3A%2F%2Fnvbugswb.nvidia.com%2FNVBugs5%2Fredir.aspx%3Furl%3D%2F3612577&data=05%7C01%7Cwilliamy%40nvidia.com%7C9d566335b54b40e4ede508da35d7bca2%7C43083d15727340c1b7db39efd9ccc17a%7C0%7C0%7C637881498070707644%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=4ySDHTxLCg5CgkCo52Mee3NMD2zgEYhMoCEZE8n%2FTlo%3D&reserved=0).

Currently 13 open issues on the [DLFW\_ARM\_Watchlist](https://nam11.safelinks.protection.outlook.com/?url=https%3A%2F%2Fnvbugswb.nvidia.com%2FNVBugs5%2FDefault.aspx%3Fdvid%3D1%23200105380%2FBUGACTION%3DDev%2520-%2520Open%2520-%2520To%2520fix%3A0%3A%3A1%3A0%3A1%26MODULENAME%3D%3A0%3A%3A0%3A0%3A2%26CUSTOMKEYWORD%3D%3A0%3A1%3A0%3A0%3A1&data=05%7C01%7Cwilliamy%40nvidia.com%7C9d566335b54b40e4ede508da35d7bca2%7C43083d15727340c1b7db39efd9ccc17a%7C0%7C0%7C637881498071020122%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=tJ%2B%2Fz8nzrQcqUwhN2UuBmZHwpZ%2B1TR7C9kC8S2UzLhg%3D&reserved=0).

[1](https://nam11.safelinks.protection.outlook.com/?url=https%3A%2F%2Fnvbugswb.nvidia.com%2FNVBugs5%2FDefault.aspx%3Fdvid%3D1%23200105380%2FBUGACTION%3DDev%2520-%2520Open%2520-%2520To%2520fix%3A0%3A%3A0%3A0%3A1%26MODULENAME%3D%3A0%3A%3A0%3A0%3A2%26CUSTOMKEYWORD%3DDLQA-Recommended-Release-Blocker%3A0%3A1%3A1%3A0%3A1&data=05%7C01%7Cwilliamy%40nvidia.com%7C9d566335b54b40e4ede508da35d7bca2%7C43083d15727340c1b7db39efd9ccc17a%7C0%7C0%7C637881498071020122%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=QOax97gb%2Ftl1nG5Dto%2BYiYx%2BUCqZHwWaDZfD5smQsFU%3D&reserved=0) DLQA-Recommended Release Blockers for training-related scenarios on Ampere Altra systems. Ampere systems are the only ones with GPUS powerful enough to consider for training scenarios.

SATA drive communication bottleneck is believed to be the source of performance variability ([3578029](https://nam11.safelinks.protection.outlook.com/?url=https%3A%2F%2Fnvbugswb.nvidia.com%2FNVBugs5%2Fredir.aspx%3Furl%3D%2F3578029&data=05%7C01%7Cwilliamy%40nvidia.com%7C9d566335b54b40e4ede508da35d7bca2%7C43083d15727340c1b7db39efd9ccc17a%7C0%7C0%7C637881498071020122%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=QlJ1mGLG6Ot4ITQPwisLYOdY2rECuREPI73FcXO%2BvZQ%3D&reserved=0)), will swap to an M.2 SSD to see if this change reduces variability ([3578029](https://nam11.safelinks.protection.outlook.com/?url=https%3A%2F%2Fnvbugswb.nvidia.com%2FNVBugs5%2Fredir.aspx%3Furl%3D%2F3578029&data=05%7C01%7Cwilliamy%40nvidia.com%7C9d566335b54b40e4ede508da35d7bca2%7C43083d15727340c1b7db39efd9ccc17a%7C0%7C0%7C637881498071020122%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=QlJ1mGLG6Ot4ITQPwisLYOdY2rECuREPI73FcXO%2BvZQ%3D&reserved=0)).

No DLQA-Recommended blockers for inference scenarios for an ARM SBSA system (Ampere, Graviton2, ROY, McCoy, or Mandalore).

Ubuntu 22.04 is being enabled in “cudnn-8.5/cuda11.7” test matrix ([3636714](https://nam11.safelinks.protection.outlook.com/?url=https%3A%2F%2Fnvbugswb.nvidia.com%2FNVBugs5%2Fredir.aspx%3Furl%3D%2F3636714&data=05%7C01%7Cwilliamy%40nvidia.com%7C9d566335b54b40e4ede508da35d7bca2%7C43083d15727340c1b7db39efd9ccc17a%7C0%7C0%7C637881498071020122%7CUnknown%7CTWFpbGZsb3d8eyJWIjoiMC4wLjAwMDAiLCJQIjoiV2luMzIiLCJBTiI6Ik1haWwiLCJXVCI6Mn0%3D%7C3000%7C%7C%7C&sdata=2xwKRZ1UsqGI2UTNoZgUqZaK0jn18R6vXLGcG4ZiOiE%3D&reserved=0))

Before Devs enable VectorCAST builds in CI (DVS\_SC), QA will keep local Jenkins automation runs to generate report.

Enabled automation to measure the code coverage on A100/V100/T4/P100 to make it easier to identify dead code.

Created the first coverage report on multiple GPU arches

I will be OOTO on April 7 and will not have access to email or to my cell phone most of this day.

I will catch up on email during the weekend and be back on line on Monday 4/11.

If you need to talk with someone before I return, please use this contact list:

The Deep Learning teams have been using ARM SBSA SW stacks (leveraging Ubuntu 2004) on Mandalore and McCoy, and you made a comment in the Deep Learning Program meeting that the intention of using the ARM SBSA stack on McCoy and Mandalore was to exclude iGPU use. Is it correct to assume the Holoscan SDK plans to always leverage DL products inside of an Ubuntu container?

I want to make sure you do not intend to ask Deep Learning to support both L4T/JetPack and ARM SBSA release packages on McCoy and Mandalore. The cost of supporting both is non-trivial and will require ~2x more hardware and engineering bandwidth to support, and the cost of supporting scenarios involving some workflow running on the iGPU simultaneously with another workflow leveraging the dGPU requires even more resources.

Raising priority to P0. DLQA needs this Luna to be fully functional or we do not have sufficient bandwidth to meet monthly DGXA100 test requirements.

If we simply close them as NAB/WNF, it was a waste of machine bandwidth and QA resource to test it (in 6 days).

Should not we evaluate these bugs affect internal/external customers when we delivered TRT-8.2.3.0, and document them in release notes (as known issue) if necessary?

Due to this SoL schedule, bug fixes on the older versions is not expected. Fixes will be picked up in the following versions aligned to Holoscan’s POR.

Thanks for your collaboration to determine why we could not get both GPUS working in these systems.  
  
I recommend bookmarking this issue in case you encounter similar issues with other systems in the future. Hopefully you will be able to cut and paste or ask other people to read a certain range of comments to help explain the actions you would like someone else to take in their role to investigate that future issue.

It sounds like you and I have to be much more explicit in messages to each other.   
Here is my very explicit request.

It is great to see the code coverage improvements are catching perf issues like bug [3505091](https://nvbugswb.nvidia.com/NVBugs5/redir.aspx?url=/3505091) much earlier than it would have been seen in the past (either by a DLFW team working on a tight MLPerf deadline, or by a customer that would want an immediate fix).

The effort to map tests to specific areas of code is going to make it much easier to determine where additional test development will likely lead to significant quality improvements as well as help identify innovation opportunities.

Many thanks to David Woon for delivering these GPUs to the IPP team

The Deep Learning QA team did not agree to give up this system. We depend on this system to have enough bandwidth to support a wide range of DL products. There simply is not enough "spare" bandwidth available in Colossus to make up for this "lost" system.

Keep up the good work!

William is working with Tinkle to get more resumes

I am just commenting to thank George for the pictures. This is one of those instances where "a picture is worth a thousand words".

给team 的改变

About Team efficiency I will improve team’s work efficiency.

I will analysis the time spend for each team members,

Make more team members role and effect from test to developer, make more automation task replace manually task. Improve automation test scripts to save test time and triage time.

And we can do more preparation before release/task upcoming, and then we can reduce test time/and test loading when the task coming.

I will re-clear the old opened bugs, make the bugs can more identify current project result.

I will unify the new bugs format and the content for each kinds bug, (such as function bugs, RFE, document bugs), and then we can automation the functional bugs open, verify to close,

Second about improve cuDNN quality

One is From my viewpoint, cuDNN library is for saving time for deep learning training and inference, but if end user developers want to use cuDNN library, they read the library api first, I think wade you also read it before, It’s so complex. I suggest we need make the cuDNN api more simply for end user, even there are much much complex implement in the code, and the gpu hardware, but it could be one the black box for user. the end user doesn’t spend much time on how to call cuDNN. They can use the library as one basic device, such as one fork, one knife, not use it very hardly and carefully, that will more friendly for end user.

second, the deep learning grow fast, and nvidia’s gpu architecture and compute capability grow very very fast too, but currently the old Kepler and maxwell gpus still supported by cuDNN, but the capability for kepler/maxwell is very low for deep learning training and inference, so is this support still necessary for cuDNN lib new version, maybe if the customer which still use the old gpus, they can use the old cuDNN version, but for new gpus,(vola/turing/ampere/hopper..) they can use new versions, that can make our library more simply.

怎么评价员工的performance

The performance evaluate is Base on the output for each engineer, such as the bugs opened, bugs fixed, and code update, and teams work efficiency improvement. And the support for the team member and the support for others team also need think about.

Key Results is does he improve cuDNN quality.

Does he improve teams work efficiency? (not only the just one person efficiency but for all team members)

怎么处理员工的冲突

If there are conflict between team member happened, first I think we need reduce the mood in the conflict(the mood in the work and discuss is very very bad), second I need very clear the question and each side’s option, and I believe team member in our team is focus on ensuring the library quality, so base on it, we can found the greatest common divisor (GCD) for the both side, base on the GCD and ensure the quality object, we can find the best method to fix the conflict.

怎么push developer/other teams

Firstly I need clearly express my requirement, why I need it, if do it, what is the benefit. If we don’t do it, what is the risk.

And for some Scenarios, I still need give the data to proof my requirement.

And if there are no feedback, I think I can email him first, if still no feedback, I will slack him directly, if still no feedback, call telephone to him, or go to his work place to talk with him directly, haha.

你相对其他的优势

First I very family with our team, our team members, our product, and the developer team, so we can start very smoothly, and work with all team member and others team very well.

Second, I also have some thinking about improve teams efficiency and cuDNN quality, and know currently status, I will do some change for our team, I will split it in several steps, and each day have little improve, but from long time see, we can get the big updated too, and I think it will reduce the reacting force and make it run smoothly.

And I can learn from the others managers, give suggestion from kevin, demy, williamg, shalia.

The last, I believe I can be one good manager, haha

你会做出些什么改变

In the previous, I like to do the tasks lone, not like to assign to the other team members, if I am one manager, I need change, and assign tasks to team members and make all team resource to achieve the objective, push related engineer and collect the status.

为什么之前没有做。

Because I’m one test developer engineer before, I should do as one good engineer, most of time I need focus on the detail tasks, the nightly/weekly test, the bugs, the release, etc, in last year, I opened more than fifty percent bugs for our team, and ramp up new team members, and developer automation scripts.