ORGANISATION EUROPÉENNE POUR LA RECHERCHE NUCLÉAIRE
EUROPEAN ORGANIZATION FOR NUCLEAR RESEARCH
Laboratoire Européen pour la Physique des Particules
European Laboratory for Particle Physics

GENÈVE, SUISSE
GENEVA, SWITZERLAND

Adresse postale / Postal address:

CERN
CH - 1211 GENÈVE 23

Téléphone / Telephone :
Direct +41 75 411 7490
General +41 75 767 6111
Electronic mail : paolo.durante@cern.ch

22 April 2021

PhD thesis assessment for Ing. Ondřej ŠUBRT

To whom it may concern,

what follows is my assessment of Ing. Ondřej Šubrt's PhD thesis, as requested.

The thesis covers the design, implementation and optimization of several vital data acquisition components of the COMPASS experiment at CERN. Ondřej's contributions to the iFDAQ system have been tested in the field over the course of several years and yielded a remarkable improvement in the reliability of the experiment's operation. The designs and algorithms are presented in detail and in accordance with standard software engineering practice. The exposition will certainly contribute to the scientific understanding of dataflow optimization problems in the context of high-throughput data acquisition systems, a topic of high contemporary relevance.

I have made a number of minor comments that you will find attached to this letter.

After taking all of the above into consideration, I believe the goal of the thesis to have been achieved and recommend it for defense.

Sincerely,

Paolo Durante,
Staff Engineer
CERN/EP-LBC
(LHCb experiment, Computing group)

Comments

pg.19: "average maximum rate": either average or maximum.

pg.20: "The Test Case" → "The Test Case 2".

pg.22: "another detectors": another is singular.

pg.22: "for the ALICE at CERN" → "for the ALICE experiment at CERN".

pg.23: "amount of collected data is steadily increasing": please quantify.

pg.23: "upgrades required due to obsolete technology": please elaborate or reference upgrade publication, if any.

pg.23: "Nowadays, ... replacement finishes.": use present progressive tense.

pg.23: "It is further divided": who? I assume "The DHC is further divided".

pg.24: The system is first described with 8 DHCmx and 8 storage nodes (1:1 ratio), the system actually installed has a 25

pg.25: What is a spillbuffer card? I assume some FPGA-based platform, but that is not specified at this point.

pg.25: At what stage of the data flow does the data format translation happen? Is it required for both stored events and monitored events? Does it affect the throughput of the system?

pg.27: "the IPC mechanisms": is plural, omit "the".

pg.27: Another question not in the list: for a distributed application, do all systems use compatible data representations? (for example regarding word endianness and, if applicable, floating point representation) DIM supports mixing little- and big-endian systems for example.

pg.28: The "most basic version of UNIX" used STREAMS as its original IPC mechanism https://en.wikipedia.org/wiki/STREAMS . Historically, this predates the implementation of pipes.

pg.28: the distinction between "shared memory" and "mapped memory" is not clear, I assume the second refers to "memory mapped files"?

pg.30: "it's" is a verb, the possessive form is "its".

pg.31: Have other industry-standard IPC libraries been considered before deciding for a rewrite? (e.g. ZeroMQ, gRPC).

pg.32: "the rest system" → "the rest of the system".

pg.32: Since the same "service" and "command" names are used in both DIM and DIALOG, and the second is meant as a replacement of the former, it would be useful to describe the differences between DIM services (and commands) and their namesakes in DIALOG.

pg.32: When the Control Server migrates to another machine, how can the clients find it? Are the spares stored in a list that is shared to all processes in advance?

pg.32: "It gives all the DIALOG functionality to an existing process just by inserting one or two lines of code" since it is only a couple lines of code, maybe an example can be shown?

pg.32: Based on these requirements, out of all the IPC primitives presented in the previous section, which ones does DIALOG use and why? I assume TCP sockets, but that is not mentioned.

pg.33: Even though the names are self-explanatory, no mention of the OutgoingThread and IncomingThread objects appears before in the document, add a forward reference to Figure 3.5. Also it is only now mentioned that messages have a header structure, but such structure is not described, add a forward reference to pg.37.

pg.33: "or kill it if its behaviour is anomalous": how can a control server kill a remote process that is not sending heartbeats? (this mechanism is also mentioned at pg.34 but not described).

pg.34: How are malicious processes prevented from connecting? Is some authentication mechanism involved?

pg.35: "the Control Server sends the message to the process, that it is not connected and probably it would like to connect.": what is the purpose of this? Is this a "reverse heartbeat" to determine that the Control Server is still running?

pg.37: Explain why using pointers can improve performance.

pg.37: The different message types, their meaning and how they are processed could maybe be better described as a table than as long text paragraphs.

pg.38: Receiver sockets are kept open until shutdown by the sender process, is there a limit to how many can be opened or can a misbehaving sender overload the system by repeatedly opening new sockets? In other words are n and m constants? What are typical values?

pg.40: Since it does not appear in the XML, where is the data format of each service and command described? In DIM, services and commands have to provide type signatures when they are registered, does DIALOG have an equivalent mechanism?

pg.42: In the GUI, messages appear to be composed of a binary header followed by a '—'-separated list of strings. This data representation (and its relationship with the format of the data on the wire) is not described however.

pg.44: "transmition" → "transmission"

pg.44/45: "The development of web application should be finished in 2018 and deployment is planned in the late 2018." "A development of the DIALOG WebSockets Daemon providing WebSockets interface is planned for year 2018": We are in 2021, are those dates still accurate?

pg.45: According to the text, the y axis of Figure 3.10 should be messages per second.

pg.51: "Moreover, Google Breakpad requires the most recent version of C++ compiler and Python interpreter. Unfortunately, such a requirement is not possible to meet in the COMPASS experiment.": since C++ is designed to be forwards-compatible, it would be relevant to describe why upgrading the compiler was not possible.

pg.53: According to the text, the system is deployed without optimizations to simplify debugging. It would be very relevant to compare the performance of the system with and without optimizations in order to quantify the cost to be paid for better observability of the system. If the performance tradeoff is negligible it should be mentioned.

pg.58: Given that this is based on Qt's signal/slot system, and given that Qt provides several types of signal/slot connections in multithreaded environments ("auto", "direct", "queued", "blocking", "unique"), it would be relevant to specify which type of connection is being used between each thread subclass instance and the debugger and why.

pg.61: This reads like memory corruption in one machine would require the reboot of all readout computers, if that is the case, explain why.

pg.65: "Recently, a stability of DAQ has become a vital precondition for a successful data taking in high-energy physics experiments.": arguably, it has always been important and not only recently.

pg.65: "malicious" implies intentional harm, unless someone was actively sabotaging the DAQ system a different word should be chosen.

pg.65: "The state with the continuously running iFDAQ is about to achieve.": confusing sentence, what is the subject and what is being achieved?

pg.65: "is used **to** initiate".

pg.67: "useless" → "unnecessary".

pg.67: "to safe" → "to save": safe is adjective, save is verb.

pg.67: "to do" → "to be done".

pg.72: "executes a DIM command": probably a DIALOG command?

pg.74: "for a detection of a new run detection" → "for a detection of a new run".

pg.74: "Apparently" → Evidently/Obviously/Naturally.

pg.78: "flows varying from 0 B to 10 kB": kB is a scalar, not a flow, I would assume the unit to be kB/s (this applies in general throughout the document, for ex. Table 7.1 at pg.135)

pg.87: Listing 6.1, 6.2, 6,8, the fibonacci function is standalone and the listing can be simplified by removing the Qt dependency altogether.

pg.96: "The key question being still remaining not answered" → "The key question still not answered".

pg.104: z and z^* are only defined in a table at pg.16 but not when introduced in the text itself.

pg.109: "well-know" → "well-known".

pg.109: "Thus, the best solution of a given problem can be found.": it is not guaranteed to find the best solution.

pg.109: "The theory of natural selection says that stronger and more powerful individuals have bigger chance to survive": it does not select for *strength*, but rather for *fitness*.

pg.109: "In general, such a stronger individual is usually faster, more intelligent or more powerful than its competitors.": same as above.

pg.111: "Mutation operator is able to return back value of element already lost.": unclear.

pg.119: "he" $\rightarrow$ "it" or "they" (use agendered singular pronouns for agent and player).

pg.120: "In order to update the value of $Q(a)$": I believe that refers to $\hat{Q}(a)$ (since $Q(a)$ is unknown and unknowable).

pg.120: "One such an efficient algorithm": "one" and "an" are redundant together, use just "an efficient algorithm".

pg.122: "Apparently" $\rightarrow$ Evidently/Obviously/Naturally.

pg.124: A step of the algorithm is sometimes denoted as k and sometimes as s, a uniform notation would be preferable. This would also avoid introducing a new y variable in 6.64, since the next state after step k is already defined on the previous page as x_{k+1}.

pg.124: Q^* is not defined (except in a table 100 pages away from this page), I can assume from previously established convention that is refers to "the optimum Q", but an explicit mention would be preferable.

pg.127: Symbol f is used to denote both flows being allocated, and the function determining state transitions, using a different letter for one of the two would be less ambiguous.

pg.132: "The learnt values in the lookup table can be used unless" $\rightarrow$ "can be reused until".

pg.137: "The C++ executions are slightly faster than the executions in Matlab, however, the difference is quite low": would be better to quantify this comparison, also at first sight this looks like a 30% difference, not quite so insignificant.

pg.138: "the implementations in both programming languages use the intern built-up ILP solving libraries" $\rightarrow$ "the built-in ILP solving libraries" (same at pg.146).

pg.139: "To conclude, ILP prepared the best possible LB for the Test Case 1": as evidenced by other implementations, there are multiple globally optimal solutions, so ILP did not find "the best" solution but one possible optimal solution.

pg.139: "larger test cases - the Test Case 2 and Test Case 3 - show how poor performance": only Test Case 1 is mentioned up to this point, preferably add a forward reference pointing to where these results can be found.

pg.141: "to demonstrate how well-designed the MDE algorithm is proposed" $\rightarrow$ "to demonstrate how well-designed the proposed MDE algorithm is".

pg.142: "Therefore, the mutual comparison reaches 73060/73060 = 100.00%": however according to table 7.8 the RL error is not always zero.

pg.142: "Unfortunately, there are no pointers in Matlab and therefore, the submatrix must be always copied to perform a particular mathematical operation": the statement is too broad to be generally valid, to the extent that Matlab operations can be (and often are) internally implemented in native languages (C, C++ or FORTRAN).

pg.146: "it is necessary to adjust slightly the ILP model for LB to be less strictly": use "strict", "strictly" is adverb.

pg.147: What does the $TolInteger$ parameter control?

pg.158: When defining which algorithms are suitable for running online at COMPASS, what is the threshold above which a solution becomes too slow to be considered for real-time use?