Development of the EU Green Public Procurement Criteria (GPP) for Data Centres

EEB comments on the 2nd draft of the JRC technical report and GPP criteria proposal

26 June 2018
Stephane Arditi, European Environmental Bureau (EEB)
with technical input by Marc-Andree Wolf, maki Consulting Gmbh

Contact

Stephane Arditi, EEB, Brussels – stephane.arditi@eeb.org

European Environmental Bureau
Europe's largest network of
environmental citizens organisations

International non-profit association –
Association internationale sans but lucratif

Boulevard de Waterloo 34, B-1000 Brussels
Tel.: +32 2 289 10 90
Email: eeb@eeb.org

EC register for interest representatives:
Identification number 06798511314-27

Introduction

The European Commission (EC) is developing EU Green Public Procurement (GPP) criteria for Data Centres (DC). In May 2018, the Joint Research Centre (JRC) of the EC published the second draft of their technical report including proposals for GPP criteria. EU GPP criteria are formulated either as Selection criteria (SC), Technical specifications (TS), Award criteria (AC) or Contract performance clauses (C). For each set of criteria there is a choice between two levels of environmental ambition: core criteria and comprehensive criteria.

Based on the discussions of JRC's proposals during the webinar of the Ad-Hoc Working Group (AHWG) in May 2018, this paper provides recommendations and comments on the proposed GPP criteria on behalf of the EEB. The EEB has consolidated this position together with its member organisations and other environmental NGOs.

Note: This paper replicates and complements comments made on the official BATIS comment spreadsheet, also circulated to JRC team.

The EEB welcomes the continuous effort to develop GPP criteria for Data Centres, including its key electronic and electrical components and also Data Centre services (all being shortened as DC in the rest of the text). We particularly appreciate that this second draft includes some of our remarks as made in our first position paper (here).

We remind that GPP criteria for Data Centres should be set with a clear higher ambition compared to existing baseline products & services and minimum legal requirements. They should be able to reflect the fast innovation of the sector and be up to date with market evolution, that's why we reiterate our idea to consider a staged approach for ambition level of some criteria and also preserve flexibility for innovation.

It is also important that GPP criteria adopt as far as possible a procurement perspective and enhance procurement decision according to what service/workload is expected from procured DC and what criteria would then be of priority significance. If this procurement perspective is deemed too difficult to capture through this document, we recommend considering the elaboration of a complementary guidance document.

The following sections should be read with this call for ambition and practicality in mind.

This paper is organised in 3 sections:

1- Priority points which we press JRC to reflect in the final version of the criteria
2- Secondary issues that we think deserve consideration and revision for the final version
3- Additional questions on which the EEB would like to get clarifications

1- Priority points to be reflected in final version of GPP criteria

1.1 - Make sure GPP criteria are future proof

Because the analysis of DC is necessarily building on past evidence, partly already many years old, and the DC industry and technology is quickly evolving, we see the need to up the ambition for GPP, to avoid that the GPP criteria - when they are finalised and put in place - are already behind the normal market and technology development. This "future-proving" of the criteria has not been implemented yet, while it is key to avoid having criteria in place that after 2 or 3 years limp behind the market instead of pulling it. **We suggest setting GPP criteria at the ambitious end of the range of possibilities, and building in a time-correction component into the criteria, i.e. auto-adjusting the criteria applicable in e.g. 2021 to be more ambitious than in 2019**. This includes - among others - criteria TS1.1.1 for server active efficiency, and others were technology advancement leads to quickly changing results for the products to be tendered.

The DC industry sees frequent innovative technology and operational developments. It is especially important for this industry, to avoid that GPP criteria block environmentally better, innovative solutions to be procurable under GPP. Only one example where such a risk exists is defining temperature & humidity ranges for server inlet temperatures or require comprehensive environment measurement systems, which could even block liquid or immersed cooling where such are not in place and not needed. The more specific, technical a criterion is (and hence the less it related to the actual environmental impact and system ICT performance), the higher is this risk. **We reiterate the need to include for each criterion a systematic check on their "innovation-flexibility"**. Either, the criterion needs to be formulated in a way that it is flexible, OR it needs to be possible for suppliers to demonstrate the environmental superiority of the solution in a different way, e.g. by submitting relevant LCA results of their solution compared to solutions on the market that meet the GPP criteria in a good way. For implementation of such "exception" cases, the Commission's PEF LCA requirements should be met to ensure appropriate comparison incl. to have a qualified, independent and external review been done on the study. **The GPP award criteria need to foresee a mechanism to consider such input for innovative solutions.**

1.2 - Adopt a system / 'package of criteria' approach rather than a mere long list of criteria

The criteria for each case (e.g. Server procurement, Cloud services, ...) need to be packaged in a way that they complement each other and jointly suitably guide identification of the most environmentally preferable offers. Still, as in the first draft, the criteria are listed and discussed mostly as a list of isolated criteria, but not as such packages, while they will form such a package automatically when put in place. An example is the PUE, which - if selected alone - is useless, but if combined with a few other criteria that control/correct its weaknesses, can add value to the criteria package. **We suggest identifying suitable packages of criteria for each procurement case, that are advised to be used jointly** - instead of a "pick and chose" by procurers - so they are comprehensive and also may overcome each other's weaknesses.
E.g. If the PUE would be one criterion *(what we rather not recommend, same as all but one commenting participants recommended during the 1st AHWG meeting*), it must be completed by criteria that consider the IT performance of the servers and the "useful work" they provide. Also, server utilisation (AC1.2.1) alone is not meaning to get an environmentally efficient data centre (as the servers might be e.g. 15 years old), but in conjunction with requiring that modern, highly performing servers are used, the criterion is very useful.

In line with this approach, we think the number of award criteria is still rather high and it will often be challenging to combine the criteria in the right way to ensure good GPP support, not affecting negatively innovation and system optimisation. **An alternative approach would avoid these issues <u>for all pro-</u>**

curements of complete DCs or DC services (i.e. not for single equipment): the overall DC use stage net energy consumption (ideally as environmental impact points) per useful work (expressed in bits/bytes sent/received with the client) combined with an EoL criterion of the DC equipment and infrastructure. This two set indicator automatically considers the ICT performance of all components and the DC as a system, incl. otherwise difficult-to-consider issues such as consolidation, virtualisation, M&E plant efficiency etc., and would be technology neutral. This would conveniently even include implicitly but reliably the ICT hardware production impacts: as the server power consumption is high per useful work at low load levels, good results for this indicator are only possible, if few, efficient servers are used at high load, meaning this also reduces the amount of ICT equipment needed and the embodied environmental burdens from production. The known limited accuracy of using the sent/received bits to express "useful work" do not too much hamper this indicator, as for a given procurement case, the data will be sufficiently comparable, moreover helped by the fact that server performance differs hugely among differently old servers. Additional aspects to consider will be to properly include energy reuse that would be deducted from the use stage energy consumption, and allocation for DC services that come from partial DCs (e.g. in co-location, cloud procurement).

1.3 - Adapt to the expected usages/services of the DC by public procurers ('workload traces')

We suggest defining 'archetypes' of expected usage scenarios (workload traces) for a limited range of the most common situations of public sector DC needs and adjust the criteria and related significance accordingly. We are aware of the uncertainties of computing and DC needs (and the always varying work profiles), but we think we do not necessarily require exact accuracy to compare among archetypal situations (unless in academics/reserach projects). Differences among servers and improvements over time are likely larger than the lack of accuracy of the archetypal workload traces that we propose here. We should not let the 'academic perfect' be the enemy of the 'practically doable good'.

1.4 – Weight the criteria according to their relevance

We call for a weighting of the different criteria (e.g by affecting more points to some) rather than considering all criteria as equally important. The award points that are assigned to different criteria need to be balanced, to reflect their relative relevance.
E.g. Server idle power is usually less relevant (as has less environmental savings potential) compared to server active efficiency. Hence, the maximum award points need to be higher for server active efficiency than for idle power.
We suggest giving quantitative guidance to procurers on the relative relevance of criteria, i.e. recommend ratios of maximum award points to be assigned/shared among the criteria. Ideally this ranking of criteria could be established per 'archetype' of expected usage, but we think a weighting is any way possible and practical.

1.5 - Question the EURECA approach to establish ideal life time of servers

The "EURECA metric" recommended for use for TS1.3.1 and CPC1.3.1 on ideal life time of servers is a very recent, theoretical publication by a single author that is based on old LCA data (as already highlighted by the JRC in the 2nd AHWG meeting) and rough assumptions (e.g: 1 GWh for producing one server, without

differentiating for example the number of processors that dominate the environmental impact of server production). Also, while the text states it would have been tested with procurers in four countries already, no evidence/report is given and considering the duration of procurement processes it appears unlikely that relevant if any experience could have been gained on it. Then, the metric ignores the highly different environmental impacts of electricity produced in different countries, that differ within the EU easily by a factor 10, rendering a common replacement rate environmentally negative in almost half of the cases and too late for almost the other half. Even if ignoring this important aspect, the replacement rate is only accurate, if the new server is much more performing than the one to be replaced, i.e. this is not about age of the server, but of the server model plus its relative performance among the servers on the market (i.e. market average server replaced with at least market average server 4 years later). However, if average servers had been used, the step to high performing models should be ensured, not just replacement again with average servers. Finally, and equally very important, to limit the determination of server replacement cycles by energy consumption alone is ignoring the fact that server production has much higher share of environmental impacts such as acidification, toxicity, resource depletion etc. than due to electricity consumption during the use phase. The currently proposed metric may hence most of the time trigger a far too frequent replacement time with subsequent environmental damages (what is good for server producers, but not for the environment or society). In addition, it neglects the potential of post warranty third party contract (third party maintenance) that have the potential to upgrade data centers when there is no sense or budget to replace existing equipment. **We suggest not using this simplified and environmentally distorted proxy measure for ideal lifetime of servers. We also call for making the case for possibility to use third party maintenance after warranty period to upgrade/maintain the performances of existing DC when it is meaningless/impossible to replace them. It could be eventually considered a maximum time period for such post warranty maintenance (e.g 3 to 4 years after warranty until 7 years age of DC) or the respect of certain conditions to ensure this is not prolonging life time at the expenses of the targeted overall environmental benefits. It has to be noted that also software can be maintained by third party (see for example** www.spinnakersupport.com or www.Origina.com).

We consider a warranty time will ensure robustness of purchased servers and prevent risks of early failures, without setting the time during which a server should be used. **We call to set the minimum warranty period in the case of purchased servers (e.g 3 years for CORE, 4 years for COMPREHENSIVE).**

1.6 – Reconsider the criteria linked to renewable energy use

The application of a criterion (AC3.1.1) rewarding use of renewable energy - with its very high improvement potential - exclusively for countries with a renewable share below 20% means that for all other countries the potential cannot be leveraged. The 20% threshold is not justified according to us, as by 2020, 20% renewable energy is a legally binding requirement, and GPP is about going beyond what is required by law. GPP criteria are to apply after 2020 and are about reducing the environmental impact of a specific product/service, also in countries with more than 20% renewable. Moreover, for countries with above 20% this can simply mean that renewables are a cost-competitive alternative, i.e. the economic hurdle would be simply lower than in other countries, so these would even be the more suitable countries to reward more renewables. **We suggest to make this criteria apply to all countries** whatever the share of renewable energy they present.

2- Secondary issues deserving consideration

2.1 - Give consideration to the impact of housing (building)

The exclusion of housing production is inappropriate: we provide references to two relevant studies (we had provided the references before, but it seems they have not been looked into) that show an environmental relevance for the housing of the same scale as the server and storage production combined. **We suggest including, <u>for new builds</u> or <u>specific major housing refurbishment</u> the building infrastructure (i.e. embodied burdens) within the scope of the analysis and GPP criteria (e.g. via building hull-related criteria, EPDs, or similar)**. A size threshold for new build and/or refurbishment could be considered for the inclusion of housing impact if deemed not relevant for minor housing work. For procurement /award criteria, there are several established EPD schemes that can be used (DGNB, BREEAM, LEED; in future the JRC-developed Level(s)), while considering the overall environmental impact results excluding the use phase of the building (as this is covered here by other criteria). The highest award points should be given to those buildings with the lowest anticipated impact (or represented at least by climate change impacts) and the ones with higher impact receiving less points. *Please see http://ieeexplore.ieee.org/document/6360435/?reload=true and* *http://maki-consulting.com/wp-content/uploads/2017/01/Ecobalance2016_Wo_presented-print.pdf*

2.2 – Set clear conditions when reuse or recycling of components of DC could be preferred

Criterion TS1.3.3 sets some specifications for reporting on reuse and recycling rate. As discussed in fact in both AHWG meetings and not disputed by any participant, the reuse of IT-perfomance critical components needs to be avoided IF THEY ARE LEADING TO INCREASED ENERGY CONSUMPTION COMPARED TO AVAILABLE MORE PERFOMING COMPONENTS, as they risk offsetting the benefits they bring with regards material resources savings. This would undermine the very core of the improvement by replacing inefficient equipment. In that case, ensuring a sound material recycling is a preferred solution. Only when clear conditions are established that re-using IT components would make sense, should it be considered. E.g when DC models are in place with no intention to replace them but only possibility is to upgrade them.

We suggest identifying possible conditions where the reuse of IT critical components could make sense (specific purchasing routes ?). In other cases, require sound material recycling of IT-performance relevant components, where technology shows a fast progress (e.g. CPU performance doubling any 18 months), i.e. the processors, expansion/graphic cards, memory. And reward reuse and further use of all other parts of the servers and of other equipment. Otherwise saying more precisions should be given with regards what is to be reported for reuse and/or recycling, or in which conditions reuse or recycling should be preferred.

Examples when reusing critical IT components could make sense:

1. Reused parts can be used as a replacement part, for repairs and/or upgrading of an even older part.
2. Sometimes technical or software requirements DEMAND to have a server in place of a certain age. Technical examples are Airplanes, Nucliar plants, Railways, the banking sector….
3. The whole IT chain consists of replacement strategies: a system of 14 years old to be replaced with a 7 year old: within the complete IT this still reduces energy

2.3 – Invite reuse of heat where it makes sense

Criteria TS2.2.1, AC2.2.1, AC2.2.2, CPC 2.2.1 are about reusing heat generated by DC. Energy reuse - while a good idea - has even more disadvantages than are currently identified in the report: it creates a structural dependency on delivering the heat, i.e. later identified; environmentally beneficial improvements in the DC with less heat produced would face the economic counter-incentive from the established heat reuse system. Moreover, the system efficiency of upgrading and transferring the low temperature heat from the DC via a heat network means that this is a low efficiency system, even if "all heat is reused". Also, many heat reuses are merely a low value waste treatment in disguise, e.g. heating poorly insulated green-houses that otherwise would not have been built. The threshold of 30% should also be avoided (as does not add value). **We suggest limiting heat reuse criteria where it makes sense that is in case of large DC (case of procuring a service to a third party or where the dissipated heat is above a certain quantitative threshold), but not imposing it to small DC rooms, where the constraints of heat reuse and lock in effect may be counterproductive**.

We would also define and exclude low value "heat reuse" from waste heat calculations. With this, we could include the request to give evidence that the purpose of the reused heat (e.g. office space heating, a greenhouse, a public swimming pool, ...) is a commercial operation/business that would be economically viable also with other, commercial heat sources. Should such a definition be practically not feasible, reuse exclusively for district heating, heating of private houses and office space should be considered for reuse accounting. (We also note that this criteria is one example, that would be automatically covered by using an award criterion based on the DC-wide net energy use, in that case would not be needed as separate criterion.)

3- Additional questions & issues to be clarified

1. While we appreciate Tables 4, 5, 6, presenting the criteria according to some public procurement routes, we still think they are presented from a criteria perspective and not a procurement route per-spective. Could it be useful to rearrange per procurement case (left column) and assign and insert the individual criteria, to obtain packages of criteria per case? This would also be the basis to derive from it the aforementioned "packages" of complementing sets of criteria per procurement case and arche-type of service/usage expectations.

2. Why is compliance with Energy Star a de facto OK for GPP (p 59)? At least should it be made clearer that it is only the latest version of Energy Star and before it is overpopulated (it is roughly estimated that after 3 years of an Energy star version, more than two thirds of models placed on the market are compliant, not making it an ambitious reference for GPP)?

3. What is 'best performing' on p 60 (= better or worse offer)? Should it be clarified and reformulated as this may seem inconsistent (more points seem to go to the DC more badly performing…).

4. Why not asking design for disassembly in CORE (not only comprehensive) p84?

5. Why emissions of hazardous substances only in COMPREHENSIVE and only for procured servers p 86? Why not asking more disclosure in COMPREHENSIVE at least? Art 33 REACH requires by law a disclosure of SVHC and candidate substances so nothing ambitious for GPP here. Would it be better to delete if only targeting SVHC, or made more challenging compared to what is anyway legally required?

6. Why so much difference in T° range for liquid cooling: Core= until 17°C and Comp until 27°C p 92?

7. Why not adopting the same approach as for other criteria to to allocate points for PUE in Core p100? Is there any logic to change the approach specifically here?

8. Why asking only environmental monitoring if no targets are set in CORE and COMPREHENSIVE p 118/119? Why no targeting certain levels of CoC achievements and differentiate between CORE and COMPREHENSIVE?

9. TS2.3.1, TS2.3.2 and followings on Environmental monitoring, cooling systems are useful where reliance on an air-based, active cooling system is needed. However these are mostly or entirely not useful where fully free cooling is possible (thinking of Iceland) or where liquid/immersed cooling is used. Such criteria would exclude superior cooling technologies. Should it be clarified that these criteria are not applicable to systems with entire free cooling or liquid/immersed cooling. Should those systems automatically be deemed to meet the TS and get the highest possible award points for these criteria?

10. Rather than information providing for service/third party with regards cooling performances on a yearly basis (CPC2.3.1, p 120), why not asking a yearly/regular improvement rate to which a bonus/penalty system could be allocated (most likely for COMPREHENSIVE)?

11. We question the usefulness of CPC 3.2.1 with regards monitoring average GHG of refrigerants emissions of cooling system compared to bid stage. Monitoring compared to what has been said at bid stage may not lead very far if not linked to an incentive mechanism such as a bonus/penalty system?

12. The table 11 setting a priority ranking of improvement area has two main weaknesses: first, it is quite confusing to have high LCA benefits (i.e. good) as "3" and high life cycle costs (i.e. bad) also as "3". In this perspective, the explanations of what 1, 2, 3 mean on page 45, bullet "d" referring to the "same scale" is unclear. Also, some values are odd e.g. for LCC - why would leasing hugely increase the LCC? Also, e.g. "Avoid over provisioning of resilience" is listed twice, with different values. Finally, the table is not consistent in its logic: "Find optimal refresh rate" has exactly the same numbers as "Ensure a high rate of utilisation of...", but is yellow, not green. Can this table be rearranged and made clearer and more consistent?

ENDS