A model to forecast data centre infrastructure costs.

To cite this article: R Vernet 2015 J. Phys.: Conf. Ser. 664 052040

View the article online for updates and enhancements.
A model to forecast data centre infrastructure costs.

R Vernet
IN2P3 Computing Center, CNRS/IN2P3, Lyon-Villeurbanne
21 avenue Pierre de Coubertin, 69627 Villeurbanne Cedex, France.
E-mail: renaud.vernet@in2p3.fr

Abstract. The computing needs in the HEP community are increasing steadily, but the current funding situation in many countries is tight. As a consequence experiments, data centres, and funding agencies have to rationalize resource usage and expenditures. CC-IN2P3 (Lyon, France) provides computing resources to many experiments including LHC, and is a major partner for astroparticle projects like LSST, CTA or Euclid. The financial cost to accommodate all these experiments is substantial and has to be planned well in advance for funding and strategic reasons. In that perspective, leveraging infrastructure expenses, electric power cost and hardware performance observed in our site over the last years, we have built a model that integrates these data and provides estimates of the investments that would be required to cater to the experiments for the mid-term future. We present how our model is built and the expenditure forecast it produces, taking into account the experiment roadmaps. We also examine the resource growth predicted by our model over the next years assuming a flat-budget scenario.

1. Introduction
The High Energy Physics community is facing a big challenge today: to satisfy the computing needs of the experiments, that are increasing steadily year after year. It is therefore of utmost importance for data centres and funding agencies to get reliable estimates about the cost to satisfy requirements of existing and future experiments.

In this paper, we provide an insight on the investments observed at a data centre, focusing on computing infrastructure and power consumption. We believe that quantifying the current costs of computing resources and measuring their trends over the last years is a way to anticipate the future budgets. This study was already carried out by CERN, for which the focus was the LHC computing [1]. The numbers obtained are a reference for LHC experiments, and we find important to compare those results with what we observe at CC-IN2P3 [2]. We studied those aspects in more detail and present our own measurements and trends.

This paper aims at providing a relation between the resource capacity provided and the corresponding yearly investment needed. This is a simple way to calculate the cost to support an experiment given its computing requirements. The first part of this paper presents the concept of modeled cost that is used for that purpose. Then we describe how we estimate the pure hardware costs observed by our site. The third part shows how we estimate the power consumption cost. Last, we present the final results, namely the investment needed for a given resource allocation plan.
2. Modeled hardware cost

The relation between a site investment, the computing capacity it provides and the evolution of the hardware price with time is not trivial. Indeed, there is always a part of a data centre budget that is dedicated to hardware replacement, another part for capacity growth, in addition to other expenditures like network backbone, batch and storage system licenses etc. In order to forecast data centre expenditures in an easy way, for what concerns hardware resources we consider in this paper the replacement and growth only. Since other investments may vary a lot from site to site, they should be calculated separately.

Providing a computing capacity has a cost, that we want to normalize to the duration it is deployed for. This comes directly from the fact that computing resources have a lifetime. In other words, for a given period of time \( dt \), the investment cost is related to the provided capacity according to Eq. 1, where \( dC \) is the investment cost of given resource for the duration \( dt \), and \( K \) is the provided (or installed) resource capacity.

\[
dC(K, t) = K(t) \cdot c_k^*(t) \cdot dt
\]

Here we introduce the variable \( c_k^* \) that binds investment and capacity, which we call the ”modeled cost”. It will be described in the next paragraphs.

When the time unit for budgets and investments is one year as it is the case at CC-IN2P3, the additional and replacement capacities are purchased once per year, so we can express Eq. 1 on a yearly basis, where \( y \) is a given year and \( C \) its respective budget (Eq. 2). The time unit can be defined differently according to the site local policy.

\[
C(K, y) = \int_{\text{year}} dC = K(y) \cdot c_k^*(y)
\]

2.1. Price evolution

It has been observed since decades that computing technology trends evolve exponentially [3], for quantities such as price per storage unit or performance per Watt, for example. One can use the Compound Annual Growth Rate (CAGR) in such exponential scenarios to quantify how a quantity drops or increases per year. In the following, we will often write \( r \) instead of CAGR for the sake of readability, both are equivalent. The CAGR for a \( c \) function is defined by Eq. 3.

\[
\text{CAGR} = r = \left( \frac{c(t + \Delta t)}{c(t)} \right)^{1/\Delta t} - 1
\]

Hence, the resource price \( c_k \) follows a profile such as \( c_k(t) = c_k(0) \cdot e^{t \times \ln(1+r)} \). In the case of a price decreasing with time, \( r \) is negative.

2.2. Hardware replacement

Computing resources in data centres are usually purchased for a given duration that may be defined by the resource warranty time \( \tau \). One may of course provide out-of-warranty hardware, but we consider here a data centre that provides capacities under warranty. In this case, in order to maintain the resource capacity over time, the hardware has to be replaced after the warranty expires. When these conditions are fulfilled,

- a constant investment is performed every year (including replacement and growth),
- the resource price \( c_k \) decreases exponentially with time,

then one can demonstrate that the model cost \( c_k^* \) follows Eq. 4.

\[
c_k^*(t) = \frac{r \cdot c_k(t)}{(1 + r)^\tau - 1}
\]
This modeled cost is a yearly cost, it allows for the fact that a resource is purchased for a certain duration ($\tau$) and fully includes the capacity to be replaced at the end of warranty. The modeled cost makes possible a simple relation between provided capacity and investment cost at a given time $t$ as shown on Eq. 2. It translates directly a capacity required by an experiment into a budget by a simple multiplication. Vice versa, it translates an available budget into a possible resource deployment roadmap.

The hypothesis of an investment constant with time is the ideal case. As a matter of fact, deviations up to $\pm 20\%$ in the budget is a reasonable window in which the modeled cost remains accurate.

From Eq. 4 one can make several observations. First the modeled cost $c_k^*$ is proportional to the actual purchase price $c_k$, which is expected. Second, when $\tau$ increases, i.e. when the resource warranty is longer, the modeled cost (which is an annual cost) naturally decreases. For example, assuming $\tau = 1$, which means all the capacity purchased at a given year has to be replaced the year after, then $c_k^* = c_k$ and the annual investment is directly proportional to the resource procurement price $c_k$ as shown in Eq. 2. Finally, when $r$ is close to zero, i.e. when the resource price does not change with time, then no growth is possible at constant budget and the modeled cost is simply $c_k/\tau$.

3. Hardware cost

3.1. CPU and disk costs

Fig. 1 shows the evolution of the CPU and disk (DAS) unitary prices observed at CC-IN2P3 for several years. These prices include computing and storage servers and part of the surrounding equipment (racks, network adapters and switches). Other hardware to operate the computing and storage services are not included. The solid coloured curves represent the procurement prices observed at various times, and the dotted coloured curves stand for exponential fits. They are referred as "Equipped Cost" on the figure. The grey curves represent the corresponding modeled costs: their unit is different from the coloured ones, since they are on a yearly basis. The last fit points show our estimate of the HS06 and TB modeled cost by mid-2014: $\sim 3 \, €$/HS06 and $\sim 50 \, €$/TB per year, respectively.

![Figure 1. Evolution of HS06 and disk TB cost. Solid: data; dotted: fit.](image)

3.2. Tape costs

The whole tape system investment includes several components such as cartridges, libraries, drives, movers, maintenance and license (HPSS in the case of CC-IN2P3). We checked that the conditions to use a modeled cost in this case are verified:
• our investments for the full tape system over the last years have been approximately constant,
• the unitary tape cartridge cost drops exponentially.

Although a significant investment was done a decade ago to buy the libraries, the recent investment profile has been rather constant since then. In the following we include nevertheless the expenses for libraries to obtain a realistic tape modeled cost.

The calculation of the tape system modeled cost is not straightforward. While only a fraction $f$ of the global tape system procurement goes for cartridges, the $r$ factor (CAGR) for the entire system capacity is equal to the one of the bare cartridges in the hypothesis of a constant investment. It means that the global price per TB of tape storage drops at the same pace as the price per TB of the bare cartridges, and this does not depend at all on $f$. As a result, Eq. 4 can be scaled by $1/f$ to get the whole tape system modeled cost (see Eq. 5), where $c_k$, $r$, and $\tau$ are respectively the cartridge unitary cost, CAGR and lifetime.

$$c_k^*(t) = \frac{r}{f} \cdot \frac{c_k(t)}{(1 + r)^\tau - 1}$$

Hence, we first fitted the cartridge unitary cost profile over the last years (Fig. 2, left plot). The obtained decrease rate $r$ was then used as a fixed parameter for Eq. 5 applied to the unitary cost of the whole tape system (Fig. 2, right plot). In this exercise, the parameter $\tau$ was set to 7 years, which is the average tape lifetime observed in our data centre. One can notice the modeled cost of tape and disk is similar nowadays, at about 50 euro per TB per year. However the cost drops faster for tape than for disk, considering the current technologies.

![Figure 2. Tape system cost evolution. Left: cartridge procurement unitary cost. Right: full tape system modeled cost. Solid: data; dotted: fit.](image)

4. Power cost
The electrical power cost of the data centre depends on two main components: the unitary power price (i.e. the cost of a kWh) and the consumption of the installed resources. The following sections show their evolution with time, including both IT and cooling consumptions of installed resources.

4.1. Unitary price
We show on Fig. 3 the evolution of the power price at CC-IN2P3 since 2004. While the average
Figure 3. Evolution of the kWh price at CC-IN2P3 since 2004, with two fit windows: total period (left) and the last three years (right).

The trend since 2004 is a yearly increase by $r = 5\%$, one can observe a steeper evolution since 2012, with $r \approx 11\%$. The future CAGR will probably stand between those two values. We will consider a yearly growth rate of $r = 5\%$ in section 5.

4.2. Resource consumption.

Fig. 4 shows the unitary power consumption for CPU and disk servers estimated at CC-IN2P3. We believe these estimations are biased before 2012 due to major evolutions in our buildings, therefore the reader should consider the points before that date as an indication of the trend, rather than a robust measurement. We fitted these profiles with an exponential function in order to forecast estimates for the coming years. For the reason mentioned above, the fit was performed on the data points after 2012.

Figure 4. Evolution of CPU and disk unitary power consumptions.

The unitary power consumption was estimated for a standard use of the resources, as observed daily in our site. It does not represent a consumption under a full utilization. As a consequence, any extrapolation to the future should be valid as long as the resource utilization remains similar. The same exercise was carried out for tape power consumption; it was found to be negligible compared to that of CPU and disk.
5. Data summary and final results

Tab. 1 summarizes the current values and evolution \((r)\) of the power cost, power consumption and hardware modeled cost. The total yearly investment \(C\) can be expressed by Eq. 6, summing up hardware and power cost estimates, where \(K\) is the installed capacity of a given resource type, \(c_k^*\) the hardware modeled cost, \(q\) the hardware unitary power consumption and \(c_q\) the power price. The typical units used for \(K\), \(c_k^*\), \(q\) and \(c_q\) are HS06 (or TB), \(\text{€/HS06/year}\) (or \(\text{€/TB/year}\)), kW/HS06 (or kW/TB) and \(\text{€/kWh}\), respectively.

\[
C(t) = K(t) \cdot (c_k^*(t) + q(t) \cdot c_q(t)) \tag{6}
\]

To estimate the future costs, we take into account:

- the future capacities to be provided over time \((K)\), \textit{i.e.} the experiment requirement or site pledged capacities,
- the quantities \(c_k^*\), \(q\) and \(c_q\) and their evolution, extrapolating their current evolution \((r)\) with time.

Thus, using Eq. 6 and Tab. 1 for each kind of resource (CPU, disk and tape), we can obtain the breakdown, by resource and by investment type, of the forecast yearly cost. As an example, we can set \(K\) to the capacities to be provided at CC-IN2P3 for LHC experiments for the next 10 years. Those capacities are obtained by increasing, year after year, our currently-installed capacity by 20%, 15% and 15% respectively for CPU, disk and tape resources, as suggested in reference [1] to stay within a constant budget. Fig. 5 shows the obtained yearly investment profiles. With this model, the calculated total investment for 2014 is quite compatible to the actual 2014 budget for LHC in our data centre. This indicates the model is working as it should. For the coming years, we expect the needed budget for such roadmap to be decreasing with time (Fig. 5, right) in spite of the fact that the final electricity bill should be increasing (Fig. 5, middle). We recall that those estimates are calculated taking into account hardware and effective power consumption.

As a consequence, assuming a constant budget over the next years, the capacities provided to LHC experiments should increase faster than predicted in reference [1]. We calculated the capacities growth rate under a constant budget hypothesis, and we obtained the results shown on Tab. 2.

6. Summary

In this paper, we have described a method to quantify the cost to provide resources to HEP experiments. This cost includes hardware, its replacement, and its power consumption. We have set up a simple relation between the provided resource capacities and the yearly investment budget, with the introduction of a useful quantity that we call the "modeled cost". Parameters
Figure 5. Expected evolution of hardware and power costs for LHC experiments at CC-IN2P3. The total cost is the sum of hardware and power costs.

| Capacity | CPU | Disk | Tape |
|----------|-----|------|------|
| $r$      | 25% | 18%  | 38%  |

Table 2. Estimated capacity yearly evolution assuming a constant budget.

Like energy consumption and hardware cost have been calculated for CC-IN2P3, and may differ from site to site. We have used the cost trends observed so far to make expenditure forecasts, taking into account the LHC needs as an example, and assuming no major change will occur in the next years on the hardware technologies we are using today for HEP. We have found that CPU, disk and tape capacities could grow by 25%, 18% and 38% respectively, with a constant budget hypothesis. These numbers should however be adjusted according to other expenses we did not (or partially) take into account, such as network, software licenses or facility upgrades. Finally, this model can be applied to experiments other that LHC, using their respective resource capacity roadmaps. It may also be applied outside the HEP context, provided that the price/performance and electrical power consumption are measured with the relevant hardware and computing workflows.

References
[1] Bird I et al. CERN-LHCC-2014-014
[2] http://cc.in2p3.fr
[3] Roser M Technological Progress http://cern.ch/go/S9Z2