Central or Distributed Energy Storage for Processors with Energy Harvesting

Erol Gelenbe, FIEEE and Elif Tuğçe Ceran
Intelligent Systems and Networks Group
Electrical & Electronic Engineering Department
Imperial College, London SW7 2AZ, UK

Abstract—We consider an interconnected distributed computer system with multiple computation centres (CC) that operate with energy harvesting to improve sustainability. The intermittent energy harvesting is matched with steady demand from the CCs using energy storage (ES), e.g. batteries. Based on energy leakage from batteries, and power losses over transmission lines, we examine whether a centralised or distributed ES system provides the solution that offers the smallest response time to a fixed workload of computer jobs using the Energy Packet Network (EPN) modelling paradigm.

Index Terms—Energy Harvesting; Distributed Computer System; Energy Packet Network

I. INTRODUCTION

The massive increase in energy consumption by ICT [1] has induced much research in using energy harvesting [2], [3], [4], [5] as a means of reducing the resulting environmental impact. Such techniques, together with optimisation [6], [7] and learning [8] can be used to judiciously store and dispatch energy in complex systems such as Cloud servers, autonomous wireless sensor nodes, and within computer chips and boards. In this context, a convenient paradigm is to consider that harvested energy, just as computer jobs or data packets, consists of discrete entities that arrive at intermittent random intervals to the system that needs electrical power to operate [9], [10], [11]. This leads to the “energy packet network” (EPN) model and related prototype implementations [12], [13], [14].

In the EPN approach the discrete representation of energy in energy packets (EPs) is accompanied by a discrete representation of data packets (DPs) and discrete units of computational work (jobs). An energy storage unit is represented as a queue whose server is the output point of the battery or energy store, while its input is the stochastic flow from an energy harvesting unit or the intermittent flow from a generator. A computational server is represented as a queue of jobs with a service station containing one or more servers (e.g. a multiprocessing computer), and a router is a queue of packets with a server that forwards DPs to other nodes in the network. The queueing formalism of G-networks [15] is a good fit for this type of system: the conjunction of a job (i.e. a unit of computational execution) and an EP, together result in the execution of the job and the “consumption” of the EP. DPs can be used to exchange commands between units, so a DP may trigger the transfer of a job to another server, or it may request the transfer of an EP towards some server that requires energy, and so on. The EPs are “ordinary customers” of the G-network where EPs are stored in queues (batteries), the external arrivals of EPs are the energy produced by sources of intermittent energy, and the G-network “triggers” are DPs that signal the requests made by consumers whose energy buffers are being emptied by jobs, which are “ordinary customers” of a different class from EPs. In addition to triggering energy packets from storage centres to computational nodes, DPs may also be used to request new jobs from other computational nodes which have a backlog of available stored energy. Since intermittent sources of energy need to be coupled with energy storage for greater efficiency, we first briefly present the EPN model, and then use it to compare a centralised with a distributed architecture for storing and dispatching energy to a set of interconnected computer systems. We show that if all stored energy is shared among all consumers, and energy losses are significant along the connections between energy storage and consuming units (i.e. the computational modules), then a centralised storage facility will be more energy efficient. Also, available data [16], [17] indicate that leakage in a larger centralised energy storage (ES) unit may smaller than the sum of leakages from multiple ESs having the same overall capacity..

Fig. 1. Model where both energy and jobs are distributed on demand between computation centres.
A. The EPN Model

Consider the EPN schematically shown in Figure 1 consisting of \(N\) computation centres (CC) \(C_1, \ldots, C_N\), and \(M\) ESs \(S_k\) replenished by intermittent renewable energy sources at rate \(\lambda_k\) in watts, \(k = 1, \ldots, M\). A job on \(C_i\) is executed in average time \(\mu_i^{-1}\) seconds with one EP of energy. This allows us to establish a relation between computational work and energy consumption; it is easy to generalise this to jobs that may require multiple EPs differently in different processors by simply having jobs at \(C_i\) return probabilistically multiple times to \(C_i\) before leaving it for another CC. Note that EPs are given in energy units (e.g. joules) while the energy flows or rates in EPs per unit time correspond to power. Each computation centre also has local energy storage that is connected to the \(S_k\). The parameter \(\eta_i\) denotes the energy loss rate by leakage at \(S_k\). After completing a work step at \(C_i\), the job either goes to some other \(C_j\) with probability \(P(i, j)\) or finishes and finishes work, hence leaves the system, with probability \(1 - \sum_{l=1}^{N} P(i, l)\). As it does its work, \(C_i\) requests energy for future work from \(S_k\) with probability \(p(i, k)\); with probability \(1 - c(k, i)\) this request is rejected, while with probability \(c(k, i)\) an EP is sent from \(S_k\) to \(C_i\), resulting in the probability \(q(k, i) = p(i, k)c(k, i)\). Note that \(p_i = \sum_{k=1}^{N} p(i, k) \leq 1\) because \(C_i\) will not necessarily request an EP for each EP it consumes. Also, energy loss will occur during energy transmission between \(S_k\) and \(C_i\), at a rate \(\delta(k, i) = \delta \neq d(k, i)\) proportional to the physical distance \(d(k, i)\) between them. Hence \(S_k\) will send not one EP, but \((1 + \delta(k, i))p(i, k)c(k, i)\) EPs/unit-time to the requesting CC to compensate for losses. Thus \(S_k\)’s total energy transmission rate is \(\sum_{i=1}^{N} (1 + \delta(k, i)) p_i h_i \mu_i c(k, i) p(k, i) \leq \Lambda_k\) where \(\Lambda_k\) is the maximum EP rate (watts) at which \(S_k\) can provide power. Furthermore, based on experimental data [18] we assume that \(C_i\) consumes energy at a rate \(\rho_i \alpha_i + \pi_i\) where \(\rho_i\) is the probability that the CC is busy processing jobs, i.e. its utilisation rate, and \(\alpha_i, \pi_i\) are constants. Applying G-Network theory [19], [20], if \(h_i\) is the probability the local energy storage at \(C_i\) is non-empty, and \(Q_k\) is the probability that \(S_k\) contains at least one EP. Assuming all job and EP arrivals are Poisson and service rates are exponential, G-network theory [15] allows us to write:

\[
\rho_i = \frac{w_i + \sum_{j=1}^{N} h_j \rho_j \mu_j P(j, i)}{\mu_i h_i}, \quad R_i = \frac{1}{1 - \rho_i} \\
h_i = \frac{\gamma_i + \rho_i \mu_i h_i \sum_{k=1}^{M} Q_k q(k, i)}{\rho_i \alpha_i + \pi_i}, \\
Q_k = \frac{\lambda_k}{\sum_{i=1}^{N} \rho_i \mu_i h_i q(k, i)(1 + \delta(k, i)) + \eta_k},
\]

where \(R_i\) is the average job response time at \(C_i\).

II. CENTRALISED OR DISTRIBUTED STORAGE

Assume now that all \(N\) CCs have identical parameters and receive identical workload, and that the local power supply in watts at the CCs \(\gamma_i\) is negligible small \(\gamma_i \ll \lambda_i\). Under these conditions we would like to compare the two systems in Figure 2, to determine whether we should have a single centralised ES storing receiving energy and distributing it to the CCs, or whether it is preferable to have \(M\) identical distributed ESs shared by all CCs. We assume that all CCs have identical processing rate \(\mu\). The ESs each receive energy at rate \(\lambda\) and have leakage at rate \(\eta\) in watts. If the ESs accept to deliver the energy request by the CCs and provide them with energy when they have it, then \(c(k, i) = 1\), so that \(\sum_{k=1}^{M} \sum_{i=1}^{N} P(i, k)c(k, i) = 1\) even when \(M = 1\). Because the CCs are identical we take \(P = \sum_{j=1}^{N} P(j, i), p_i = p\). For both the centralised and distributed case we have:

\[
\rho = \frac{w + \rho h P}{\mu h}, \quad h = \frac{\gamma + \rho h Qp}{\alpha \pi + \pi}.
\]

As a consequence we have the following results:

Lemma 1 Both for the system that uses a single centralised ES or \(M\) distributed ESs, as long as all of the CCs are identical with identical workload \(w\) then:

\[
\rho h = \frac{w}{\mu(1 - P)}, \quad h = \frac{\gamma + w[Qp - \alpha]}{\pi(1 - P)},
\]

which only depends on the individual arrival rate of jobs \(w\) to each CC, the job service rate \(\mu\), and \(p, P, \pi, \alpha\).

Lemma 2 Since jobs visit on average \(1/(1 - P)\) CCs before completion, with \(N\) identical CCs the average response time for jobs in the system is \(R^* = (1/\mu)[(1 - P)(1 - \rho)]^{-1}\). \(R^*\) is an increasing function of \(\rho\). By Lemma 1 \(\rho\) is a decreasing
function of \( h \), and \( h \) is an increasing function of \( Q \), it follows that as \( Q \) increases, \( R^* \) decreases.

Now assume that for the centralised case, all CCs are at the same distance \( d = 1 \) of the ES; the probability that the centralised ES unit contains at least one EP is:

\[
Q_c = \frac{\lambda_c}{\lambda_c + \eta_c} = \frac{\lambda_c}{Np\frac{u}{T-p}(1 + \delta) + \eta_c},
\]

while for each distributed ES it is:

\[
Q_d = \frac{\lambda_d}{\lambda_d + \eta_d} = \frac{M\lambda_d}{Np\frac{w}{T-p}(1 + \delta_d) + M\eta_d},
\]

where we assume that all CCs make energy requests to all ESs equally, and \( \delta_d \) is the average energy transmission loss rate from ESs and CCs in the decentralised organisation. Thus:

\[
\frac{Q_c}{Q_d} = \frac{\lambda_c}{M\lambda_d} \frac{Np\frac{w}{T-p}(1 + \delta_d) + M\eta_d}{\lambda_c + \eta_c} = \frac{\lambda_c}{\lambda_c + \eta_c} = \frac{Q_c}{Q_d}.
\]

This proves the following result.

**Theorem** If the total harvested power supply in the centralised and distributed systems are identical, \( \lambda_c = M\lambda_d \):
- If the transmission losses are the same in both cases \( \delta = \delta_d \), and the centralised system leakage is less than the overall storage leakage rate of the distributed system \( \eta_c < M\eta_d \), then \( Q_c > Q_d \) and \( R^*_c \leq R^*_d \). Hence at equal power the centralised system delivers lower response time.
- More generally for an equal amount of harvested power in both cases, \( R^*_c \leq R^*_d \) if and only if:

\[
\delta_d - \delta \geq \frac{\eta_c - M\eta_d}{Np\frac{u}{T-p}},
\]

which only depends on the leakage and transmission losses, the number \( M \) of ESs and \( N \) of CCs, on the workload, and the probability \( p \) that the CCs make a request for energy to the ESs after processing each job.

**III. CONCLUSIONS**

We study the choice of centralised or distributed energy storage in multi-computer system, when energy flows intermittently to ESs from energy harvesters, computer jobs arrive intermittently to CCs which need energy from the ESs, and computer jobs circulate among CCs till completion. Assuming identical total energy flows and computer job flows into the system, we show that the average job response times will depend on ES energy leakage and transmission loss parameters, and make optimum choices between a large or several smaller (in total equivalent) ESs.

**Acknowledgement**

Work supported by UK EPSRC Grant No. EP/K017330/1 on “Energy harvesting Communication netwoRks: OPtimization and demonstrate” to Imperial College within ERANET ECROPS.

**REFERENCES**