femto-st SCIENCES & TECHNOLOGIES

Energy Consumption Optimization of Parallel Applications with Iterations using CPU Frequency Scaling

PhD Dissertation Defense

Ahmed Badri Muslim Fanfakh

Under the supervision of:
Raphaël COUTURIER and Jean-Claude CHARR

UBFC - FEMTO-ST - DISC Dept. - AND Team

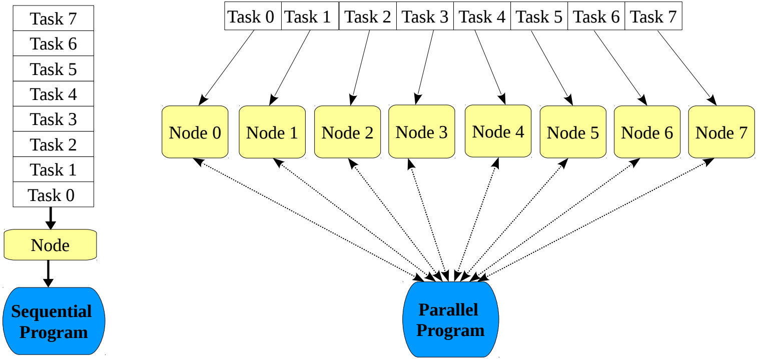
17 October 2016

UBFC UNIVERSITÉ BOURGOGNE FRANCHE-COMTÉ · UFC UNIVERSITÉ DE FRANCHE-COMTÉ · cnrs · ensmm · utbm université de technologie Belfort-Montbéliard

Outline

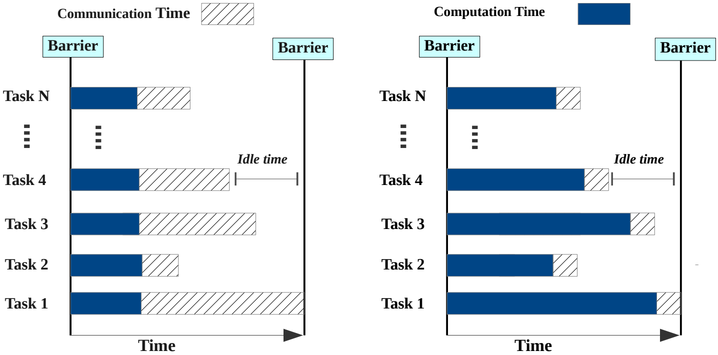
1. Introduction and Problem definition
2. Motivations
3. Energy optimization of a homogeneous platform
4. Energy optimization of a heterogeneous platform
5. Energy optimization of asynchronous applications
6. Conclusions and Perspectives

Definition of parallel computing



Sequential Computing **Parallel Computing**

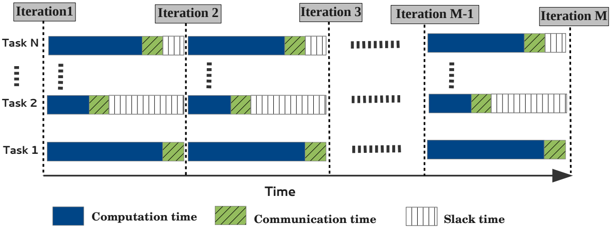
Execution of synchronous parallel tasks

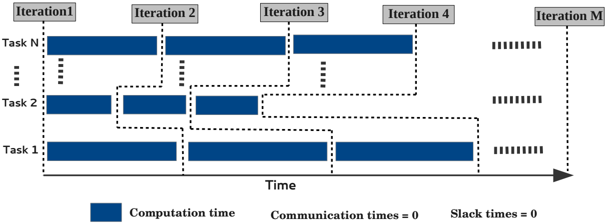


(a) Synchronous imbalanced communications

(b) Synchronous imbalanced computations

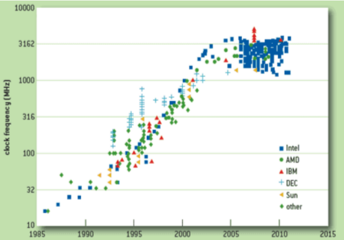
Synchronous and asynchronous iterative methods





Approaches to get more computing power

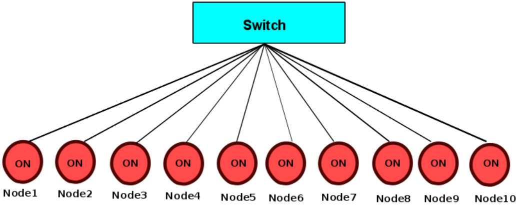
1) **Increase the frequency of a processor.**
(limited due to overheating)



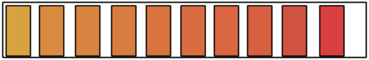
2) **Increase the number of computing units.**
The supercomputer Tianhe-2 has more than 3 million cores and consumes around 17.8 megawatts.

Techniques for energy consumption reduction

1) Switch-off idle nodes method

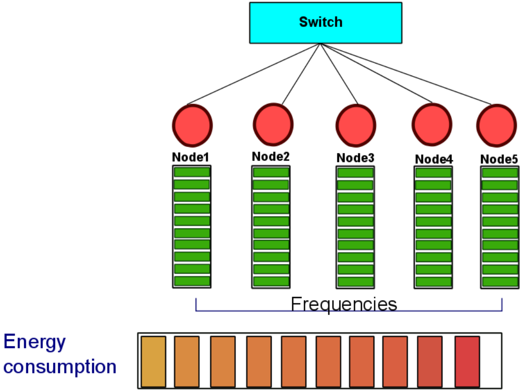


Energy consumption



Techniques for energy consumption reduction

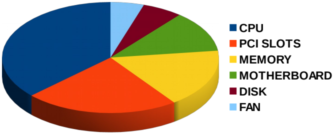
2) Dynamic Voltage and Frequency Scaling (DVFS)



Motivations

Why we used the DVFS method:

- The CPU is the component that consumes the highest amount of energy in a node [1].



- DVFS reduces the energy consumption while keeping all the nodes working.
- It has a very small overhead compared to switching-off the idle nodes.

Challenge and Objective

Challenge: DVFS is used to reduce the energy consumption, but it also degrades the performance of the CPU.

Objective: Applying the DVFS to minimize the energy consumption while maintaining the performance of the parallel application.

[1] Fan, X., Weber, W., and Barroso, L. A. 2007. Power provisioning for a warehouse-sized computer.

The first contribution

Energy optimization of a parallel application with iterations running over a homogeneous platform

Objectives

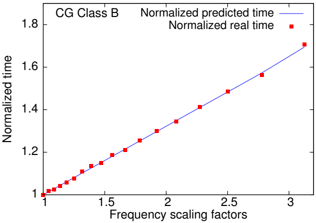
- Studying the effect of the frequency scaling on the **energy consumption and performance** of parallel applications with iterations.
- Discovering the **energy-performance trade-off relation** when changing the frequency of the processor.
- Proposing an algorithm for selecting the scaling factor that produces **the good trade-off** between the energy consumption and the performance.
- Comparing the proposed algorithm to existing methods.

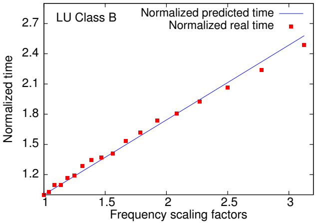
Performance evaluation of MPI programs

The frequency scaling factor is the ratio between the maximum and the new frequency, $S = \frac{F_{max}}{F_{new}}$.

Execution time prediction model

$$T_{new} = T_{MaxCompOld} \cdot S + T_{MinCommOld}$$





The maximum normalized error for CG=0.0073 **(the smallest)** and LU=0.031 **(the worst)**.

Energy model for a homogeneous platform

The power consumed by a processor is divided into two power metrics: the dynamic (P_d) and the static (P_s) powers.

$$P_d = \alpha \cdot CL \cdot V^2 \cdot F \qquad (1)$$

Where:

α: switching activity.

CL: load capacitance [F].

V: the supply voltage [V].

F: operational frequency [Hz].

$$P_s = V \cdot N_{trans} \cdot K_{design} \cdot I_{Leak} \qquad (2)$$

Where:

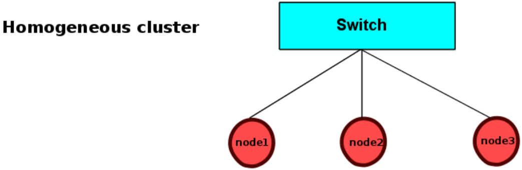
V: the supply voltage [V].

N_{trans}: number of transistors.

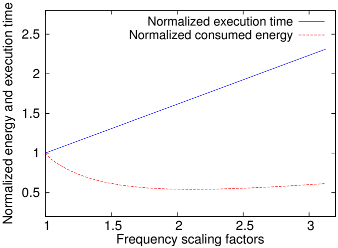
K_{design}: design dependent parameter.

I_{leak}: technology dependent parameter [A].

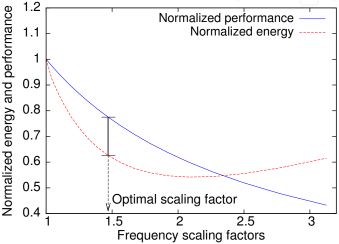
Energy model for a homogeneous platform



Performance and energy reduction trade-off



(c) Real relation.



(d) Converted relation.

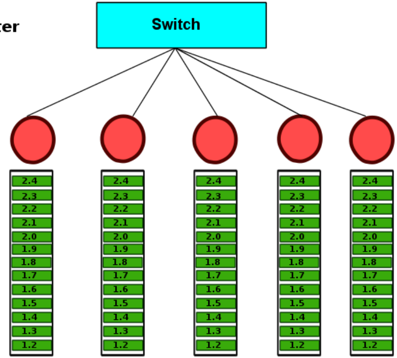
Where: $\textit{Performance} = \textit{execution time}^{-1}$

Our objective function

$$\mathbf{\textit{MaxDist}} = \max_{j=1,2,\ldots,F}\big(\overbrace{P_{Norm}(S_j)}^{\textit{Maximize}} - \overbrace{E_{Norm}(S_j)}^{\textit{Minimize}}\big)$$

Scaling factor selection algorithm

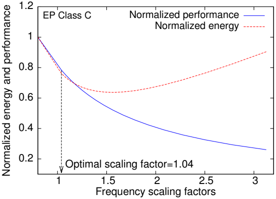
Homogeneous cluster

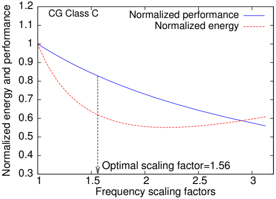


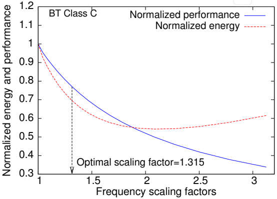
Experiment over SimGrid

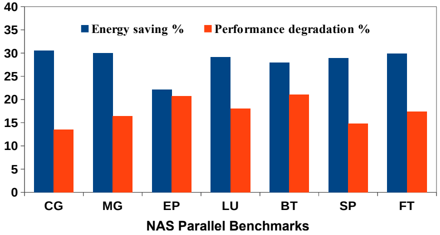
- The experiments were executed on the simulator SimGrid/SMPI v3.10.
- The proposed algorithm was applied to the NAS parallel benchmarks.
- Each node in the cluster has 18 frequency values from **2.5**GHz to **800**MHz.
- The proposed algorithm was evaluated over the A, B and C classes of the benchmarks using 4, 8 or 9 and 16 nodes respectively.
- $P_d = 20W$, $P_s = 4W$.

Experimental results





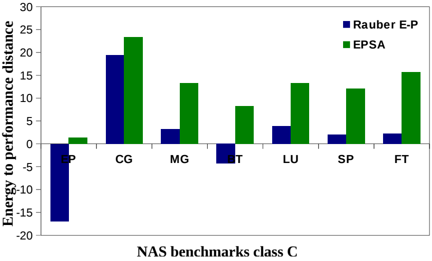




Results comparison

Rauber and Rünger's scaling factor [2]

$$S_{opt} = \sqrt[3]{\frac{2}{N} \cdot \frac{P_{dyn}}{P_{static}} \cdot \left(1 + \sum_{i=2}^{N} \frac{T_i^3}{T_1^3}\right)}$$



[2] Thomas Rauber and Gudula Rünger. Analytical modeling and simulation of the energy consumption of independent tasks. In Proceedings of the Winter Simulation Conference, 2012.

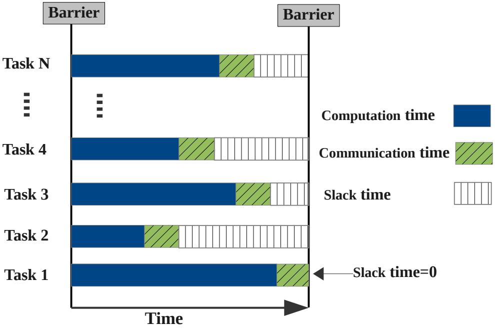
The second contribution

Energy optimization of a parallel application with iterations running over a Heterogeneous platform

Objectives

- Proposing new energy and performance models for message passing applications with iterations running over a heterogeneous platform (cluster or Grid).
- Studying the effect of the scaling factor S on both the energy consumption and the performance of message passing iterative applications.
- Computing the vector of scaling factors $(S_1, S_2, ..., S_n)$ producing the good trade-off between the energy consumption and the performance.

The execution time model

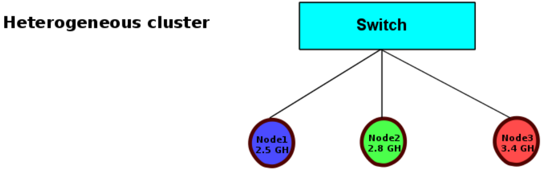


The execution time prediction model

$$T_{new} = \max_{i=1,2,\ldots,N}(TcpOld_i \cdot S_i) + \min_{i=1,2,\ldots,N}(Tcm_i) \quad (3)$$

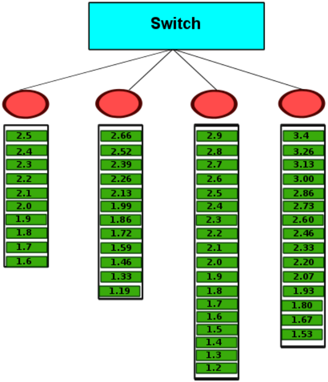
Where: $Tcm = communication\ times + slack\ times$

The energy model for heterogeneous cluster



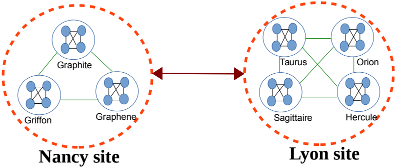
The scaling algorithm for heter. cluster

Heterogeneous cluster

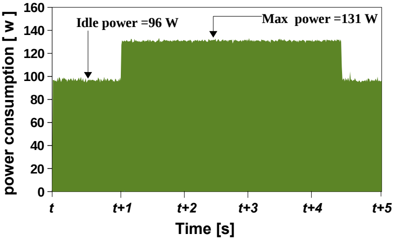


Experiments over Grid'5000

The experiments were conducted using three clusters distributed over one or two sites.

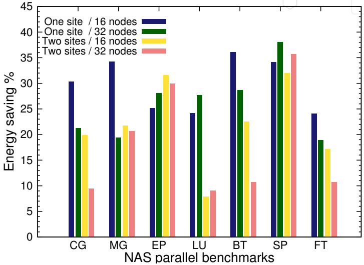


Grid'5000 power measurement tools were used.

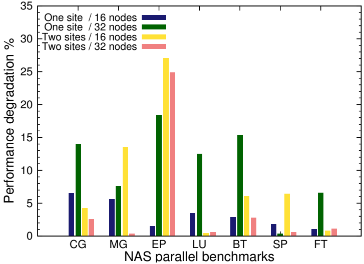


Experiments over Grid'5000

The average energy saving = 30%

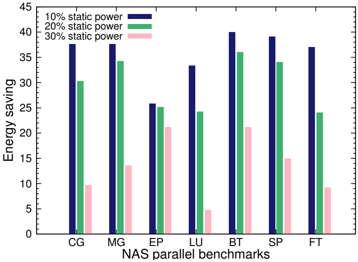


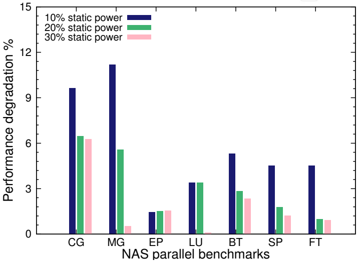
The average performance degradation = 3.2%



femto-st SCIENCES & TECHNOLOGIES

The results of the three power scenarios



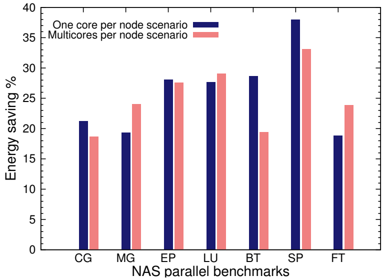


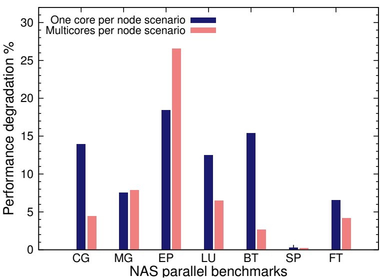
Selected frequency scaling factors

30%Pstatic scenario	1	1	1	1	1	1	1.99	1.99	1.99	1.99	1.99	1.99	1.66	1.66	1.66	1.66
20%Pstatic scenario	1.25	1.25	1.25	1.25	1.25	1.25	1.99	1.99	1.99	1.99	1.99	1.99	1.66	1.66	1.66	1.66
10%Pstatic scenario	1.25	1.25	1.25	1.25	1.25	1.25	2.49	2.49	2.49	2.49	2.49	2.49	1.66	1.66	1.66	1.66

femto-st SCIENCES & TECHNOLOGIES

One core and Multi-cores per node results

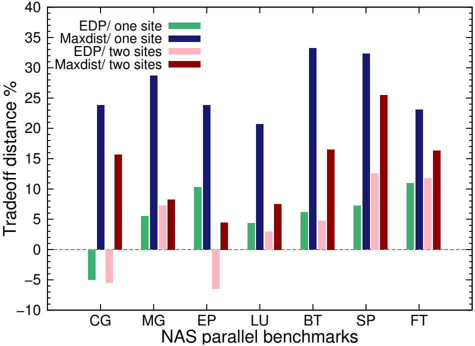




Using multi-cores per node scenario decreases the computations to communications ratio.

Comparing the objective function to EDP

EDP is the product between the energy consumption and the delay [3].



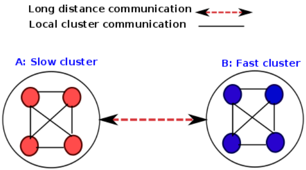
[3] Spiliopoulos et al, Green governors: A framework for continuously adaptive dvfs, in International Green Computing Conference and Workshops (IGCC), 2011.

The third contribution

Energy optimization of asynchronous iterative message passing applications

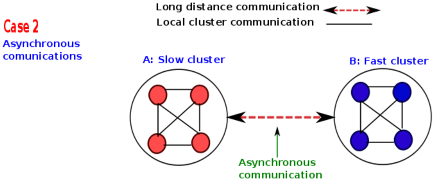
Problem definition

The execution of a synchronous parallel iterative application over a grid



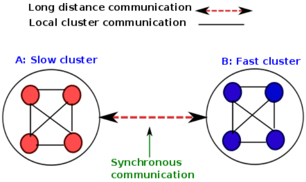
Problem definition

The execution of an asynchronous parallel iterative application over a grid



Solution

Using asynchronous communications with DVFS



The performance and the energy models

Synchronous Applications

Performance Model

Tnew = Execution time of the slowest task

Energy Consumption Model

$E = \Sigma$ Dynamic-energy(Tcp_i)+
Σ Static-energy(Tmax)

→

Asynchronous Applications

Performance Model

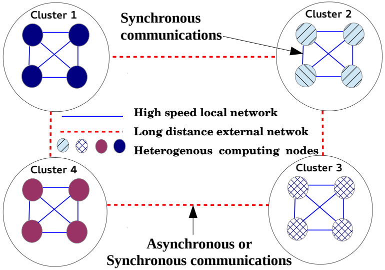
Tnew = The average of the execution time of all tasks

Energy Consumption Model

$E = \Sigma$ Dynamic-energy(Tcp_i)+
Σ Static-energy(Tcp_i)

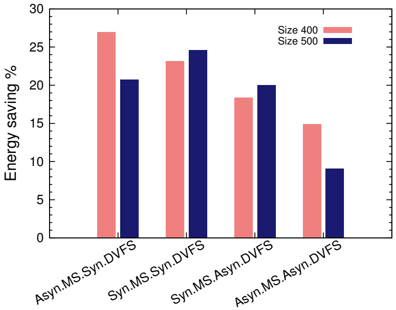
The experiments

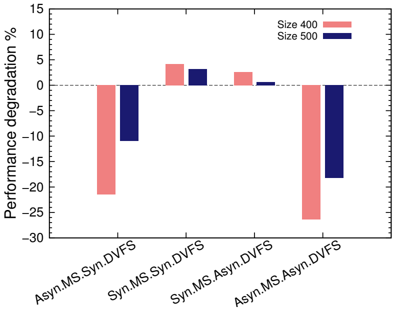
- The architecture of the grid:



- Applying the proposed algorithm to the asynchronous iterative message passing multi-splitting method.
- Evaluating the application over the simulator and Grid'5000.

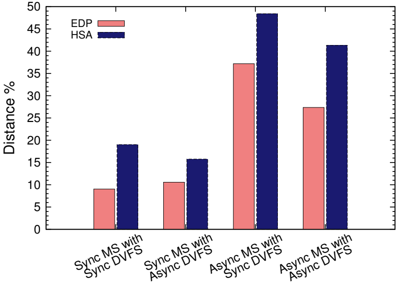
The Grid'5000 results

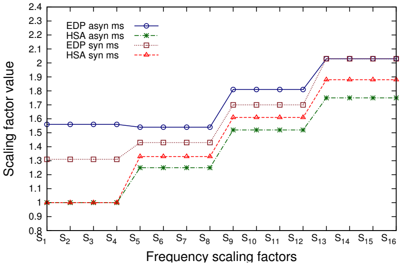




The average energy saving = 26.93%, the average speed-up = 21.48%

The comparison results





Conclusions

- Three new energy consumption and performance models were proposed for synchronous or asynchronous parallel applications with iterations running over homogeneous and heterogeneous clusters or grids.
- A new objective function to optimize both the energy consumption and the performance was proposed.
- New online frequency selecting algorithms for clusters and grids were developed.
- The proposed algorithms were applied to the NAS parallel benchmarks and the Multi-splitting method.
- The proposed algorithms were evaluated over the SimGrid simulator and over the Grid'5000 testbed.
- All the proposed methods were compared to either Rauber and Rünger's method or to the EDP objective function.

Publications

Journal Articles

[1] Ahmed Fanfakh, Jean-Claude Charr, Raphaël Couturier, Arnaud Giersch. Optimizing the energy consumption of message passing applications with iterations executed over grids. *Journal of Computational Science*, 2016.

[2] Ahmed Fanfakh, Jean-Claude Charr, Raphaël Couturier, Arnaud Giersch. Energy Consumption Reduction for Asynchronous Message Passing Applications. *Journal of Supercomputing*, 2016, (Accepted with minor revisions)

Conference Articles

[1] Jean-Claude Charr, Raphaël Couturier, Ahmed Fanfakh, Arnaud Giersch. Dynamic Frequency Scaling for Energy Consumption Reduction in Distributed MPI Programs. *ISPA 2014*, pp. 225-230. IEEE Computer Society, Milan, Italy (2014).

[2] Jean-Claude Charr, Raphaël Couturier, Ahmed Fanfakh, Arnaud Giersch. Energy Consumption Reduction with DVFS for Message Passing Iterative Applications on Heterogeneous Architectures. *The* 16^{th} *PDSEC*. pp. 922-931. IEEE Computer Society, INDIA (2015).

[3] Ahmed Fanfakh, Jean-Claude Charr, Raphaël Couturier, Arnaud Giersch. CPUs Energy Consumption Reduction for Asynchronous Parallel Methods Running over Grids. *The* 19^{th} *CSE conference*. IEEE Computer Society, Paris (2016).

Perspectives

- The proposed algorithms should take into consideration the variability between some iterations.
- The proposed algorithms should be applied to other message passing methods with iterations in order to see how they adapt to the characteristics of these methods.
- The proposed algorithms for heterogeneous platforms should be applied to heterogeneous platforms composed of CPUs and GPUs.
- Comparing the results returned by the energy models to the values given by real instruments that measure the energy consumptions of CPUs during the execution time.
- Considering the power consumed by the other devices in the node such as the memory and the hard drive in the energy consumption model.

Fin

Thank you for your attention

Questions?