

Application of GPUs to Industrial Inspection/Measurement Systems

Takeshi Saruwatari, Yuichiro Hikida, Takuya Yasuda
Research & Development Center, Dainippon Screen Mfg. Co., Ltd.



1. Background

- GPUs are now widely used for applications that involve high performance computing technologies in computational science, big data analytics and cloud computing services.
- Industrial inspection/measurement systems, our products, also employ these technologies to handle large volume data such as very high resolution images over GB size.
- Thus we intended that GPUs would have a potential to give advantage to our systems.

2. Motivation

- For industrial inspection/measurement systems, the required performance is too high for S/W based on CPUs to provide in many cases.
- To achieve the requirements, we have often developed new dedicated H/W based on FPGAs.
- H/W can provide much higher performance than S/W, but costs much more to develop and support (see Fig.1).
- GPUs make it possible to obtain H/W-comparable performance by S/W processing if they adapt to the target application.
- If we can satisfy the requirements by S/W based on GPUs, we will reduce the costs.

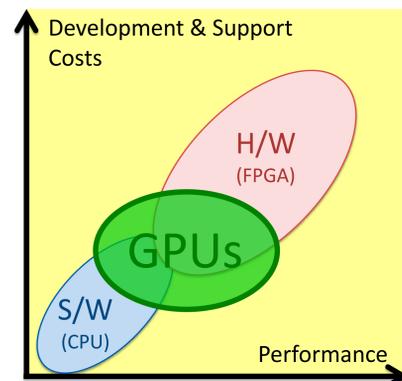


Fig.1. Tradeoff between Performance and Costs.

3. Application to Wafer AOI

3.1. What is Wafer AOI ?

- Automated optical inspection (AOI) is an inspection system by processing optical images of the specimen surface acquired with cameras.
- Our Wafer AOI system "ZI-series" targets a semiconductor wafer (disk-shaped substrate) and finds out defects of integrated circuits formed on it.



Fig.2. Appearance of ZI-series.

3.2. Motivation & Goal

Motivation :

- The previous system of ZI-series has utilized a dedicated H/W based on four FPGAs for image processing.
- But the H/W does not have enough flexibility of customization to promptly reflect user requirements.
- Thus we planned to renew the image processing system by GPU-based S/W instead of the H/W to improve flexibility (see Fig.3).

Goal :

- To implement the same inspection algorithms as that of the H/W on GPUs.
- To achieve the performance to inspect a wafer within 60 seconds with the GPU system (not slower than the H/W system).
 - Over 540MB/sec of image processing speed required.

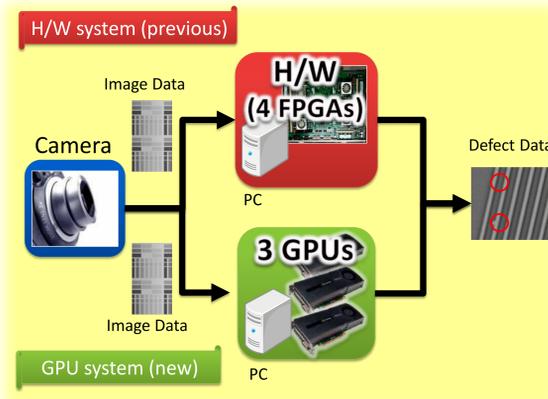


Fig.3. Difference in Configurations between the H/W and the GPU system.

3.3. Inspection Procedure

- The system acquires the images of the entire wafer surface by reciprocal scanning with a line sensor camera (see Fig.4).
- The inspection process runs parallelly with the input comparing each "die" to neighboring ones and extracting defective portions on it.
- A die is a block of circuit and several hundreds to thousands of dies that have the same form are arranged in grid on a wafer (see Fig.5).

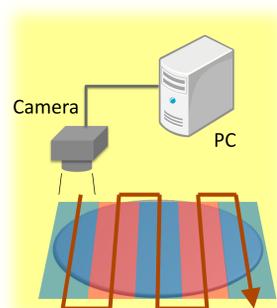


Fig.4. Scanning Sequence.

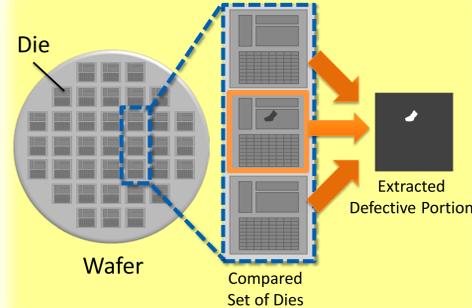


Fig.5. Inspection Method.

3.4. Performance Result

Fig.6 contrasts measured inspection time for a wafer on the H/W and the GPU system under two conditions. The "standard" condition corresponds to an inspection setting with moderate load and the "high-load" to one with particularly high load.

- For the standard, the GPU system finished the inspection at the almost same speed as the H/W one. Since the inspection time includes time for mechanical scanning motions, this is almost the minimum restricted by the scanning speed.
- For the high-load, the GPU system finished the inspection faster than the H/W one. That indicates the GPU system exceeds the H/W one in processing capacity.

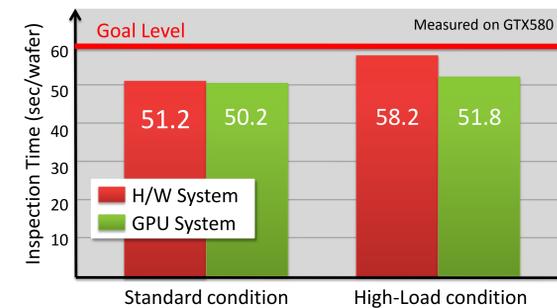


Fig.6. Comparison in Inspection Time between the H/W and the GPU System.

In conclusion, we have achieved the goal performance with GPU system.

3.5. Cost Result

Fig.7 contrasts several kinds of costs on the H/W and the GPU system expressed as relative values with the H/W one being 100 percent.

- The development costs, estimated in the actual man-hours to implement the standard inspection functions, were reduced down to about 80% less than that of the H/W one.
- The customization costs, estimated in the actual man-hours to implement an additional inspection function as an example, were reduced down to about 75 percent less than that of the H/W one.
- The parts costs, estimated in the money to produce or purchase H/W, GPUs and PC, of the GPU system (on Tesla C2075 x 3) were reduced down to about eight percent lower than that of the H/W one.
- Moreover, we have achieved the goal performance with even two GPUs of Quadro K5000, so the latest costs were reduced down to about 25 percent lower.

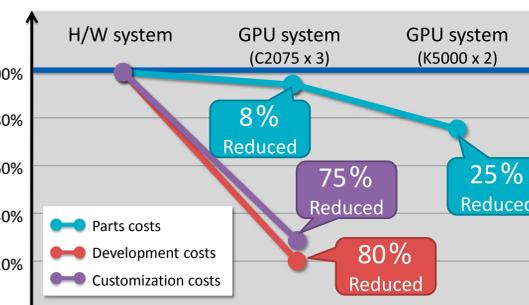


Fig.7. Comparison in Costs between the H/W and the GPU System.

Industrial products tend to have continual function additions reflecting user requirements over their life cycle. As shown in Fig.8, the total cost advantage of GPU system is expected to increase in the future.

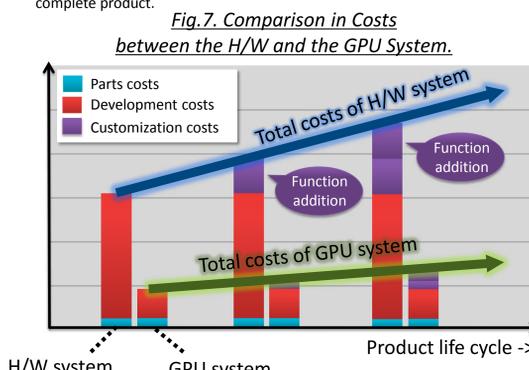
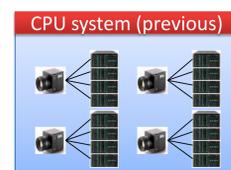


Fig.8. Expected Cost Advantage of the GPU System against the H/W over the Life Cycle.

4. Other Applications

A GPU-based PC cluster system



4 PCs per camera,
Up to 16 PCs with 4 cams.



1 PC with 1 GPU per camera,
Up to 4 PCs with 4 cams.

GPUs enhance the performance per PC to reduce the number of PCs.
-> Lower parts costs.
-> Lower maintenance costs.
-> Lower power consumption.

A GPU-H/W hybrid system



GPUs collaborate with H/W to scale-down H/W.
-> Quick product release (Low development costs).
-> High flexibility (Low support costs).

5. Summary

- We demonstrated significantly high performance especially in cost-effectiveness of GPUs through an actual application in our Wafer AOI system, wherein we replaced H/W with GPUs for image processing.
- As for that case, we obtained the performance that exceeded the H/W based on four FPGAs with two or three GPUs, and reduced costs approximately 80 percent in development costs, 75 percent in customization and eight or 25 percent in parts by comparison with the H/W system.
- Considering that continual customizations last over the product life cycle, and also the growth of GPUs may make us possible to take the lower priced model or the smaller number of ones to satisfy the requirements, the cost advantage will increase in total in the future.
- Despite the fact that GPUs cannot take the place of H/W in all cases, they are worth using to reduce costs in such ways to assist CPUs in a PC cluster system or collaborate with H/W in a hybrid system.