EINDHOVEN
UNIVERSITY OF
TECHNOLOGY

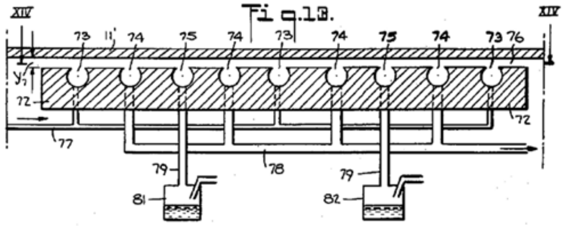
Tuorial session: Scaling up Spatial ALD

Paul Poodt – p.w.g.poodt@tue.nl

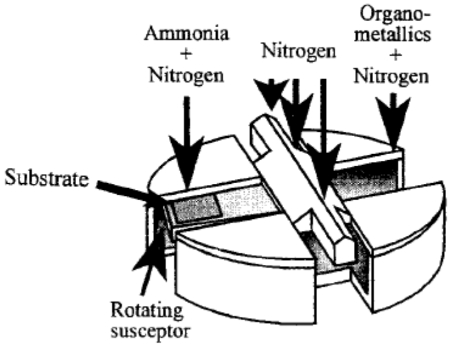
Department of Applied Physics

A brief history of Spatial Atomic Layer Deposition

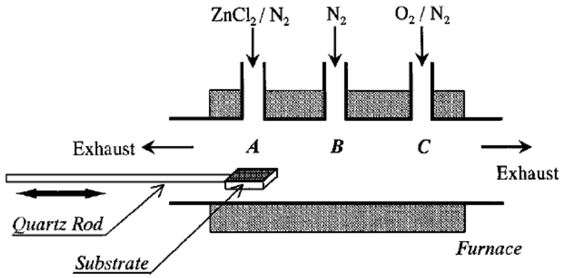
- Spatial ALD was already described in the first ALD patents by Suntola and Antson in the ‘70s. Since then it was forgotten and re-invented several times



Suntola and Antson, US patent 4,058,430 (15 Nov. 1977)



Bedair at al, J. Cryst. Growth 178 (1997) 32



K. Kaiya, et al, J. Mater. Sci. Lett. **19**, 2089 (2000).

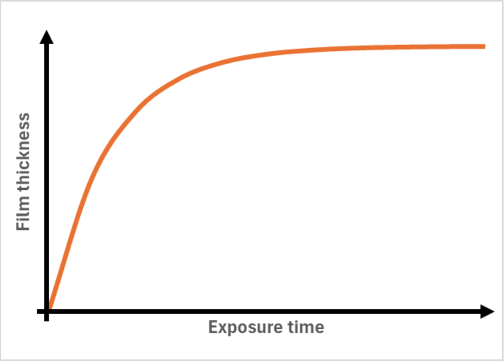
- In the late 2000’s, it was “re-invented” at various locations
  - E.g. Jusung (KOR), Eastman Kodak (US), ASM International (NL), Lotus Applied Technology (US), TNO (NL) and others
- At the 2011 ALD conference in Cambridge (US), the name “Spatial ALD” was officially introduced
  - Poodt, Cameron, Dickey, George, Kuznetsov, Parsons, Roozeboom, Sundaram & Vermeer, J. Vac. Sci. Technol. A 30 (2012) 010802

The 1st most asked question about Spatial ALD

"How fast?"

Key parameters in Atomic Layer Deposition

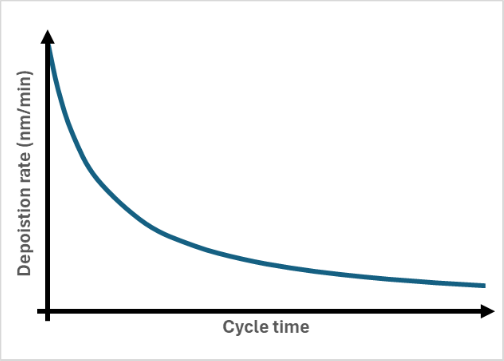
Growth per cycle (GPC)



$$GPC = \frac{nm}{cycle}$$

- GPC of Al_2O_3 ≈ **0.1 nm/cycle**

Deposition rate (R)



$$R = \frac{GPC}{cycle\ time}$$

- GPC of 0.1 nm, 5 s cycle time: ***R*** **= 1.2 nm/min**

The growth per cycle ≠ deposition rate!

Key parameters in Atomic Layer Deposition

- A more relevant parameter is **Throughput:** products produced per unit time
  - E.g. wafers, plates, cells, kg's of powder….

$$\boxed{\boldsymbol{Throughput = \frac{wafers}{hour} = \frac{n}{\frac{h}{R} + t_o}}}$$

n = batch size

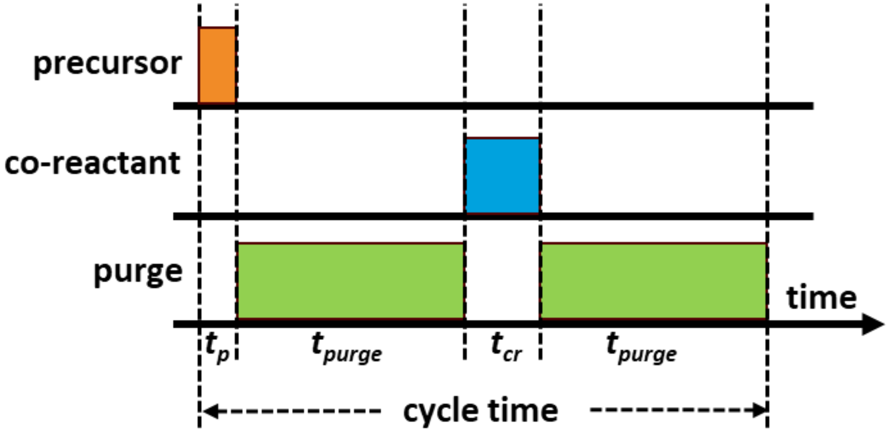
h = film thickness

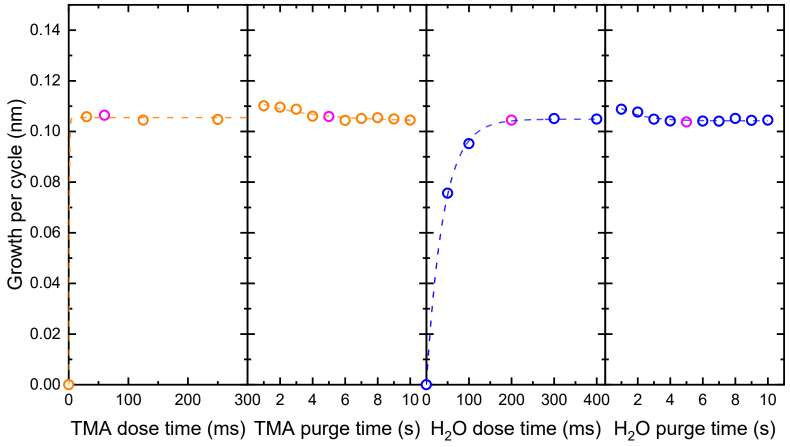
R = deposition rate

t_o = overhead time; idle time, time required for loading/unloading, heating/cooling, etc

Time scales in Atomic Layer Deposition

The cycle time





A typical ALD cycle contains:

1. A precursor- and a co-reactant **exposure step**
   - Takes ~100’s of ms

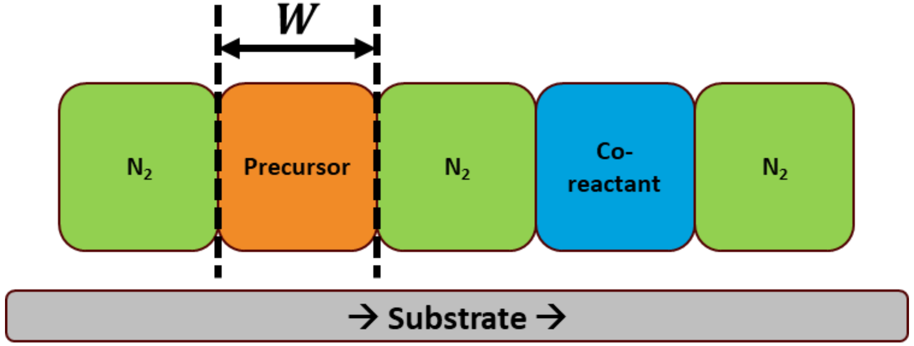
2. **Purge steps** after the precursor/co-reactant exposures
   - Takes several seconds

- The timing of the individual steps are optimized and controlled through **opening and closing valves**
- All these steps add to the **total cycle time**

Time scales in Spatial Atomic Layer Deposition

The cycle time

In Spatial ALD, the timing of the individual steps is determined by the Spatial ALD head design and substrate speed



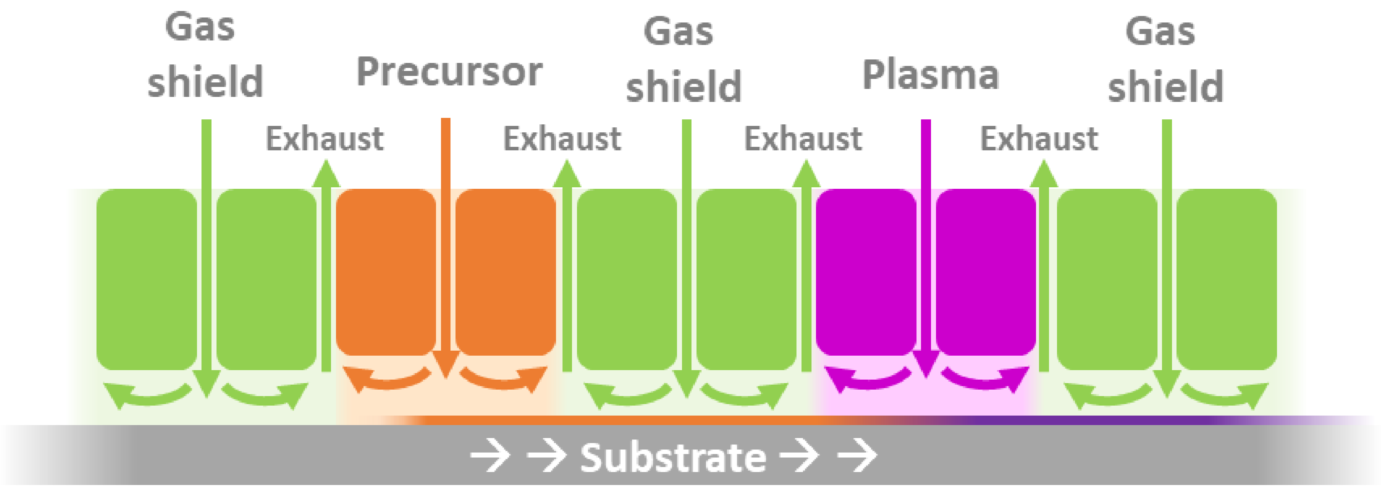
$$\boxed{t = \frac{W}{v_s}}$$

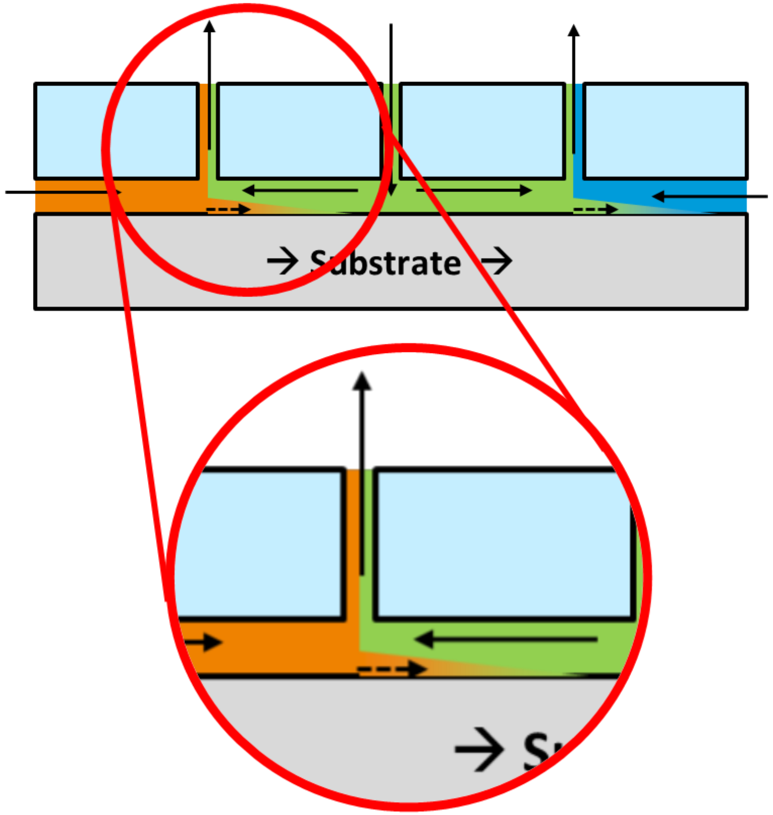
Similar to conventional ALD, we can define precursor/co-reactant exposure times and “purge” times, adding up to the Spatial ALD cycle time

Time scales in Spatial Atomic Layer Deposition

The purge time

In Spatial ALD, there is no reactor chamber to purge, but **gas shields to separate the precursor and co-reactant**



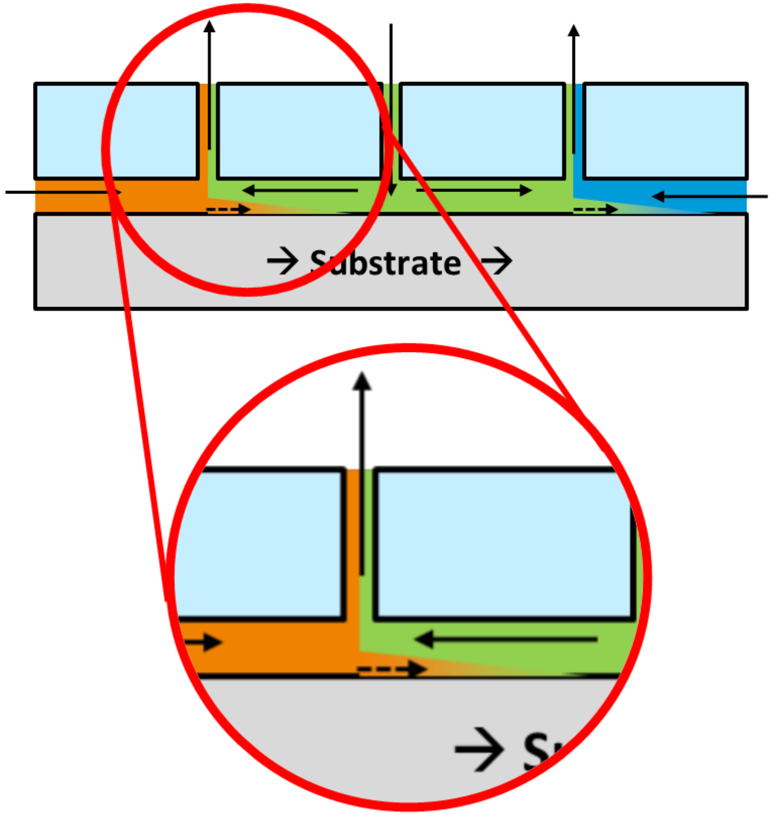


Time scales in Spatial Atomic Layer Deposition

The purge time

In Spatial ALD, there is no reactor chamber to purge, but **gas shields to separate the precursor and co-reactant**

- Substrate moves from precursor to co-reactant: precursor is **dragged along**
- An N_2 flow is used to "**push back**" the drag flow to avoid mixing
- Modelling shows: a few mm of N_2 shield is enough, corresponding to **a few ms "purge time"**
- (In practice, shield widths ≈ precursor slot widths)



Time scales in Spatial Atomic Layer Deposition

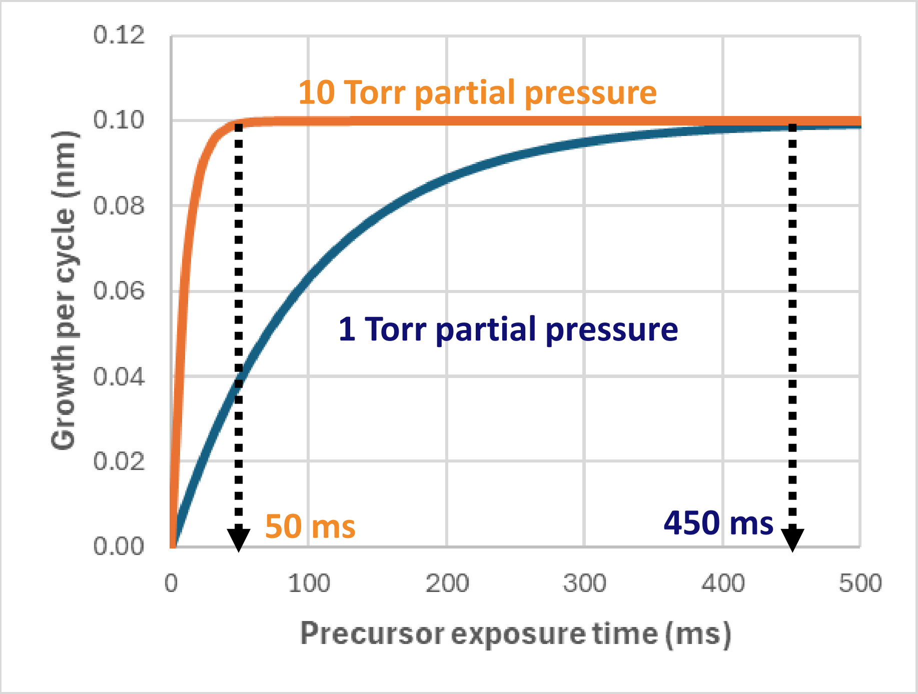
The precursor exposure time

The precursor coverage θ during exposure is given by

$$\boxed{\boldsymbol{\theta = 1 - \exp(-k\,p\,t_{\mathrm{exp}})}}$$

k: reaction rate coefficient; p: precursor partial pressure

So the time to reach saturation is given by $\boxed{\boldsymbol{t_{exp} \propto \frac{1}{p}}}$



Atmospheric pressure Spatial ALD: high partial pressure can be used to minimize exposure time

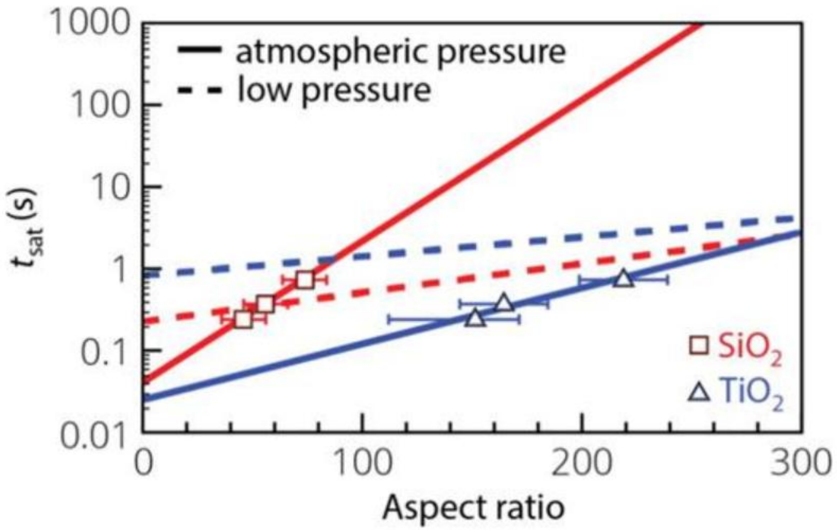
Time scales in Spatial Atomic Layer Deposition

The co-reactant exposure time

Also the co-reactant exposure time can be reduced by increasing partial pressure

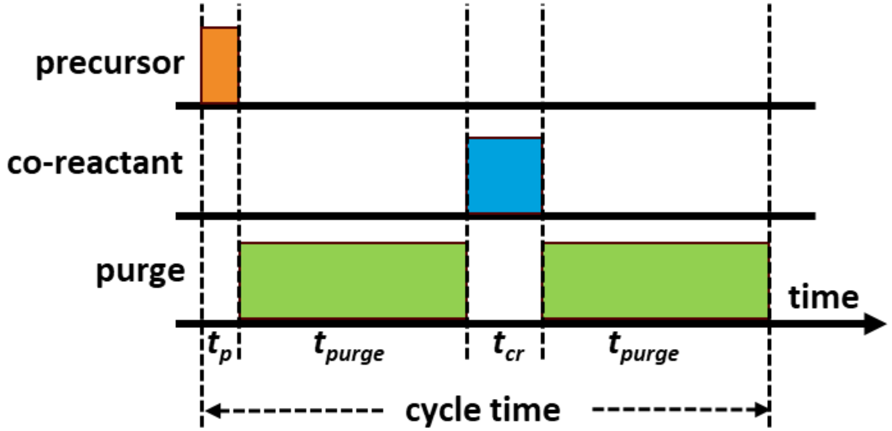
Special case: **atmospheric pressure plasma enhanced Spatial ALD**

See presentation by Mike van de Poll at 10:55 – 11:45



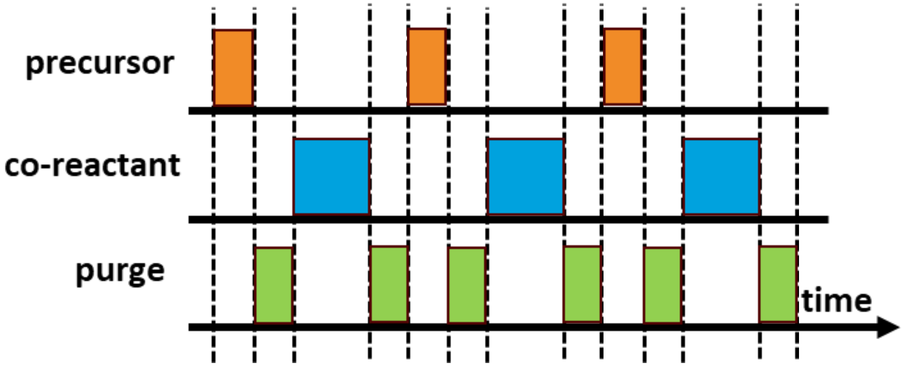
The cycle time in Spatial ALD

In **conventional, low pressure ALD**, the cycle time is dominated by the purge time



Typical cycle times are in the 4-10 s range

In **Spatial ALD**, the purge- and exposure times can be reduced to 10’s-100’s of ms



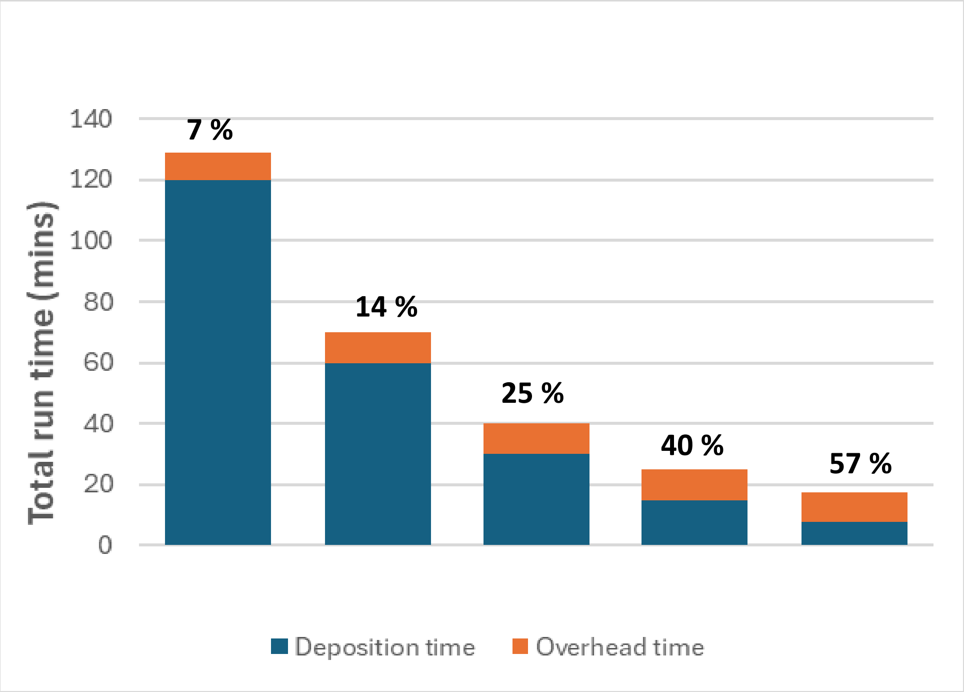
Cycle times < 1s are possible

Overhead time

- A deposition run takes more time than just the deposition
- e.g. loading & unloading, heating & cooling, pumping & venting, coffee-breaks etc.

$$\boxed{Throughput = \frac{n}{\frac{h}{R} + t_o}}$$

- Overhead time can have a big impact on throughput if it is long compared to deposition time



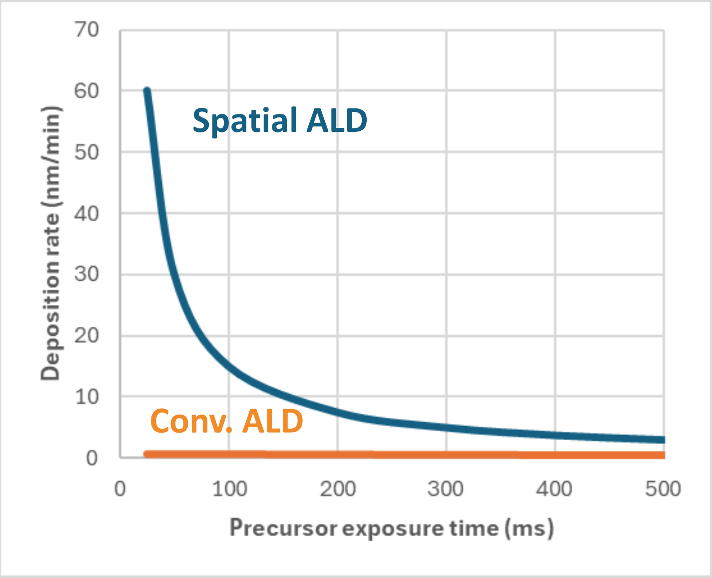
Throughput examples

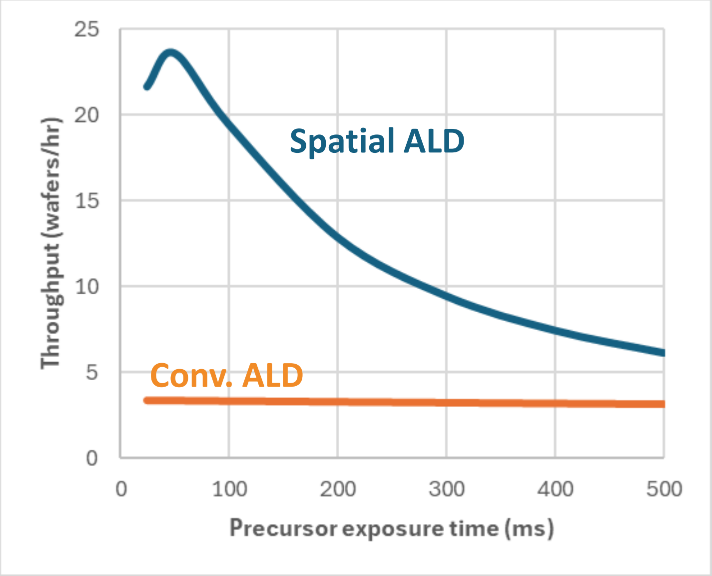
What does all this mean for throughput?

- Example: single wafer reactor
    - 6 inch wafer, 10 nm thickness, GPC = 0.1 nm/cycle
    - Spatial ALD: Exposure times = purge times
    - Conventional ALD: purge times of 5 s each
    - 120 s loading/unloading; all manual

- Deposition rate: **5-50x** higher; Throughput: **3-7x** higher

- The difference between deposition rate and throughput is because of the overhead time



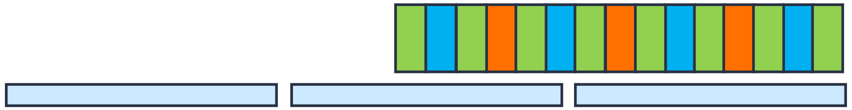


How to maximize throughput

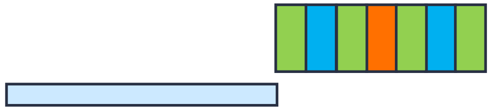
Ways to increase throughput up to 100’s wafers/hr:

1. Increase cycles per passage
2. Use “batch mode”
3. Continuous movement

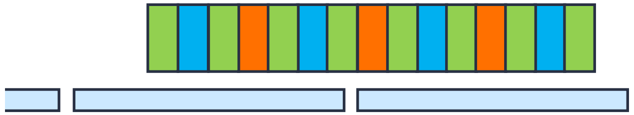
Multiple cycles per passage



Multiple wafers (“batch”)



One cycle per passage
Single wafer



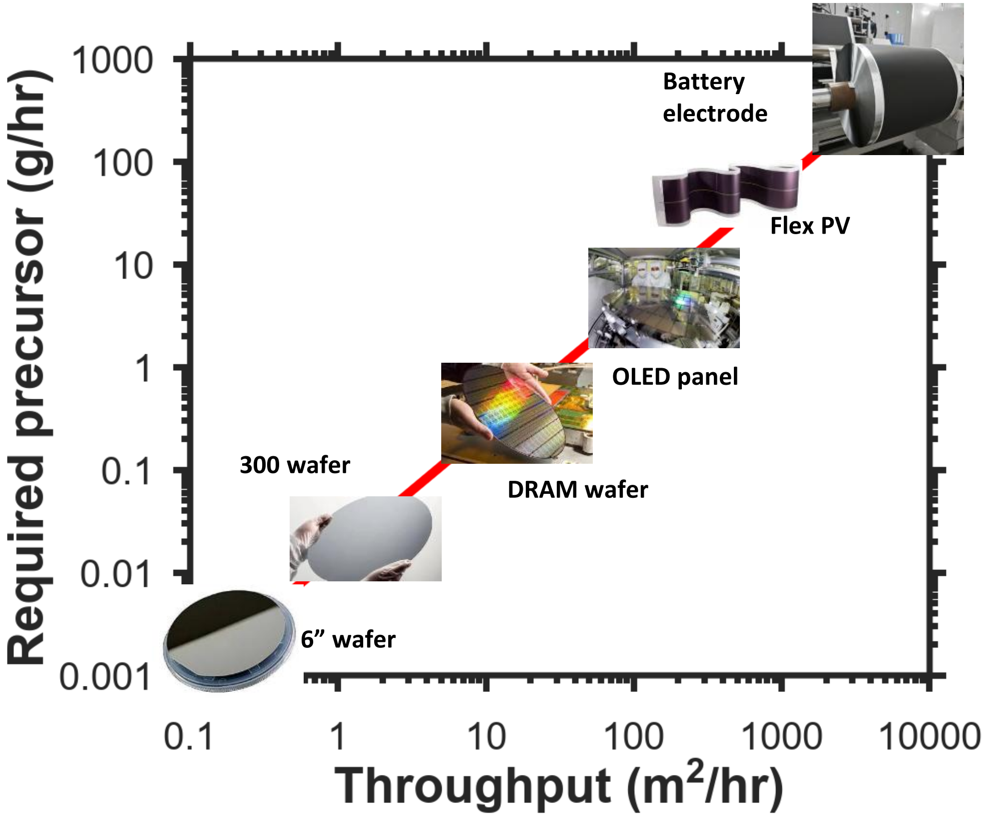
Continuous movement
(e.g. roll-to-roll)

How much precursor do we need?

- If we increase throughput, more precursor needs to be supplied

- Maximum precursor flow from a bubbler is limited by its vapor pressure:
  - TMA: ~25 g/hr
  - TDMASn: ~10 g/hr
  - $MeCpPtMe_3$: ~1 g/hr

- For low vapor pressure precursors: risk that throughput is **limited by precursor supply**



Case study

Let’s consider passivation of a graphite anode in a Li-ion battery using TMA

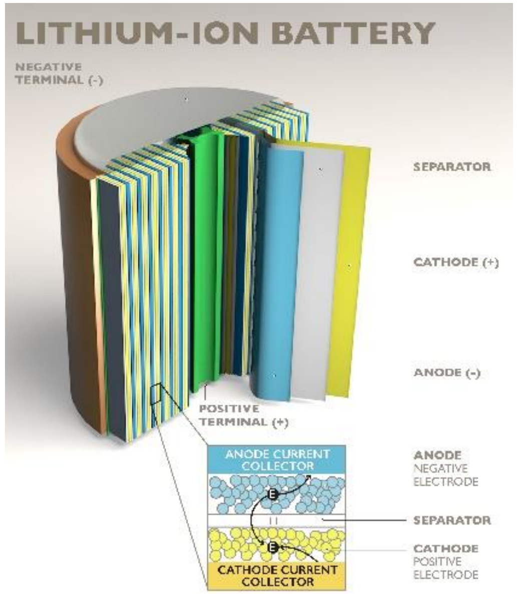
Assumptions

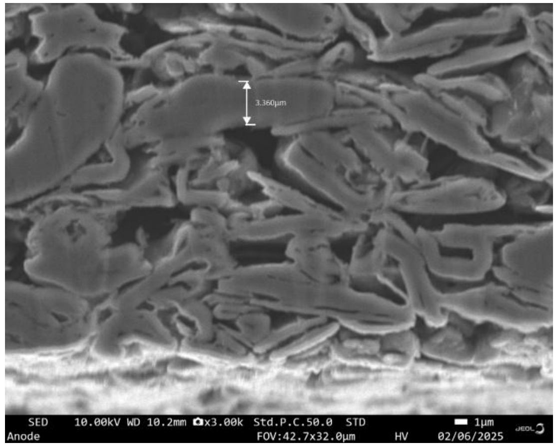
- Cathode thickness: 40 µm, effective surface area ~25 m^2/m^2
- Cathode foil roll width: 1 m; Manufacturing throughput: 90 m/min

For this throughput (2250 m^2/min!), the estimated required TMA flow is ~ **100 g/hr** per ALD cycle

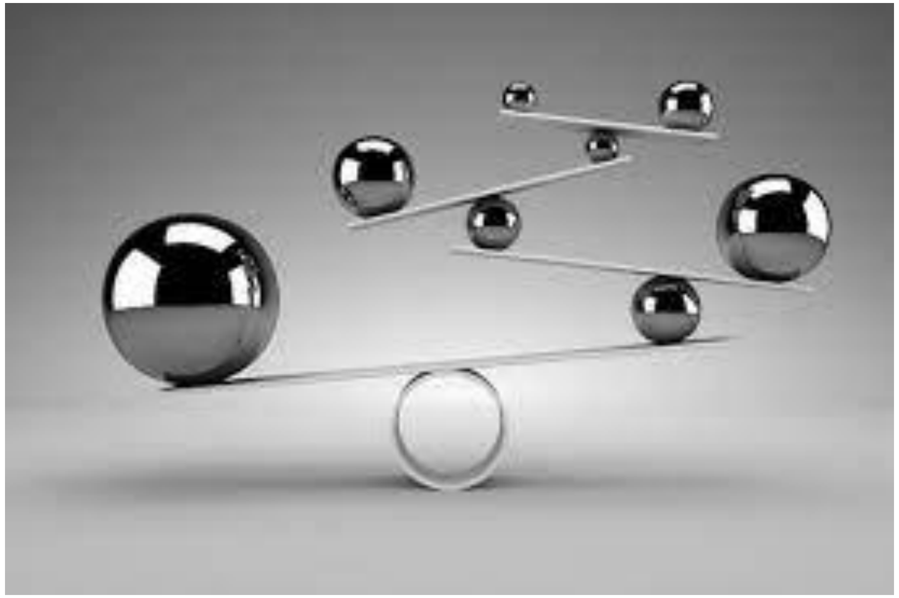
- i.e. for 10 cycles in parallel, we need to supply **1 kg TMA per hour!**

For these really high-throughput applications, precursor supply is one of the main engineering challenges





Summary



Scaling up Spatial ALD is a balancing act

- Precursor flow and partial pressure, substrate speed and ALD head design, substrate size and effective area, ...

The speed and throughput of Spatial ALD can be 10-100x higher than conventional ALD

- Short purge- and exposure times result in high deposition rates
- Optimized reactor design can minimize overhead times

For large effective surface area substrates, very large precursor flows are required

- This can be a limiting factor for low vapor pressure precursors