Cloud security
Virtualisering
Serverless security
AI security
IoT security
Hardware security

Carsten Jørgensen
Department of Computer Science

DIKU 24. september 2021

Old School vs. New World

En cloud…

Hvad skal man tænke på?

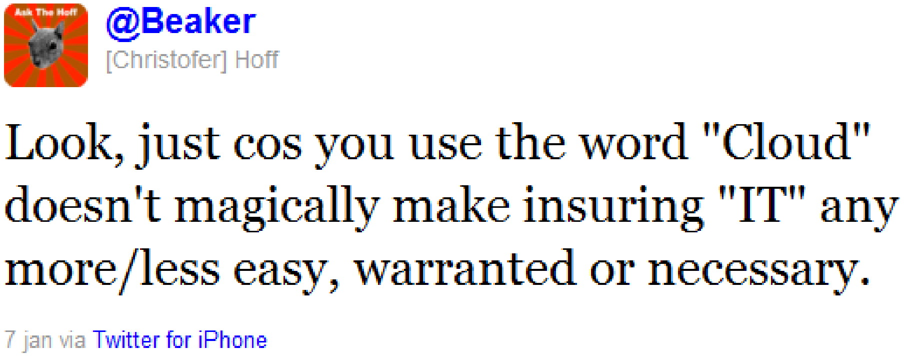
”Vi overvejer en cloud-løsning – er det sikkert?”

Eller mange gange:
”Vi har købt en cloud-løsning – er det for resten sikkert?

”Cloud” er ikke automatisk ”sikkert”

IT bliver ikke ”sikkert” på magisk vis, bare fordi man kalder noget ”cloud”

Men det bliver heller ikke usikkert



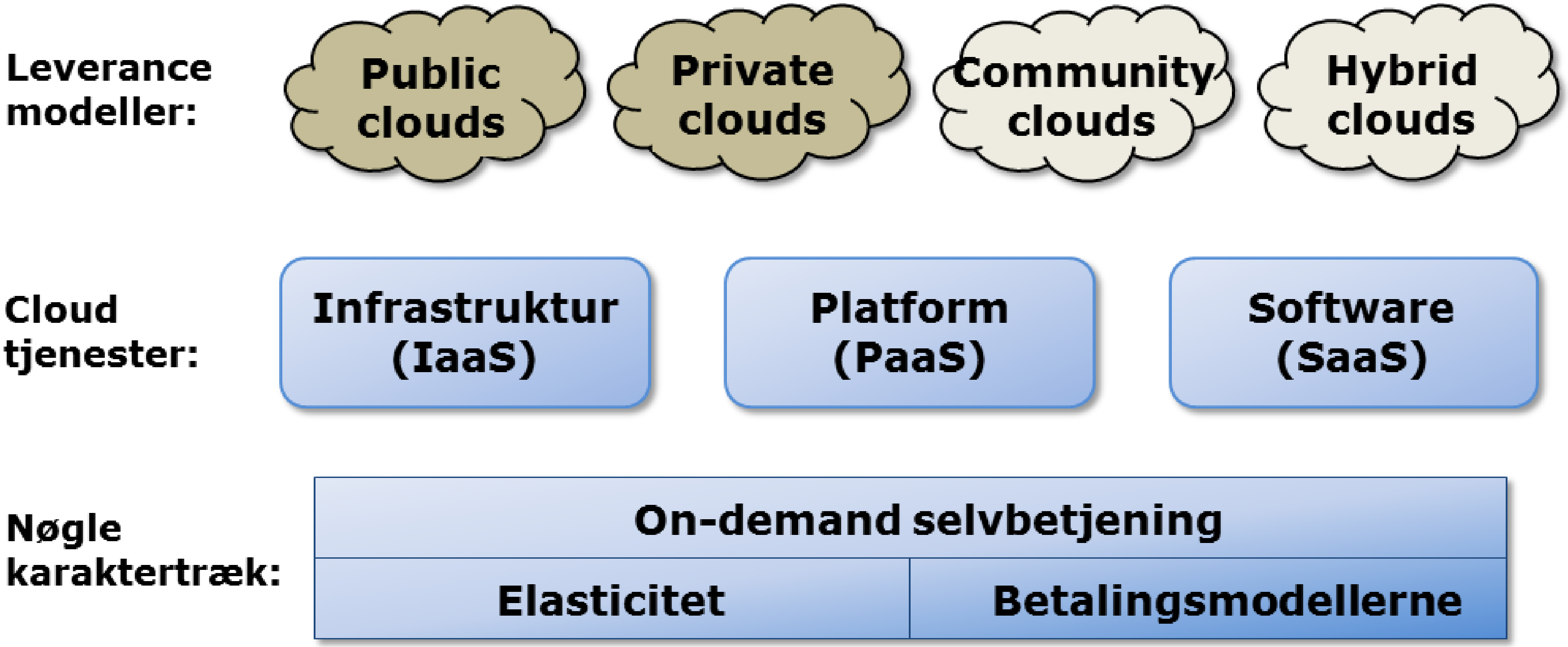
@Beaker
[Christofer] Hoff

Look, just cos you use the word "Cloud" doesn't magically make insuring "IT" any more/less easy, warranted or necessary.

7 jan via Twitter for iPhone

Hvad er cloud computing

Cloud Computing er en drifts- og leverancemodel



Leverance modeller:

Public clouds

Private clouds

Community clouds

Hybrid clouds

Cloud tjenester:

Infrastruktur (IaaS)

Platform (PaaS)

Software (SaaS)

Nøgle karaktertræk:

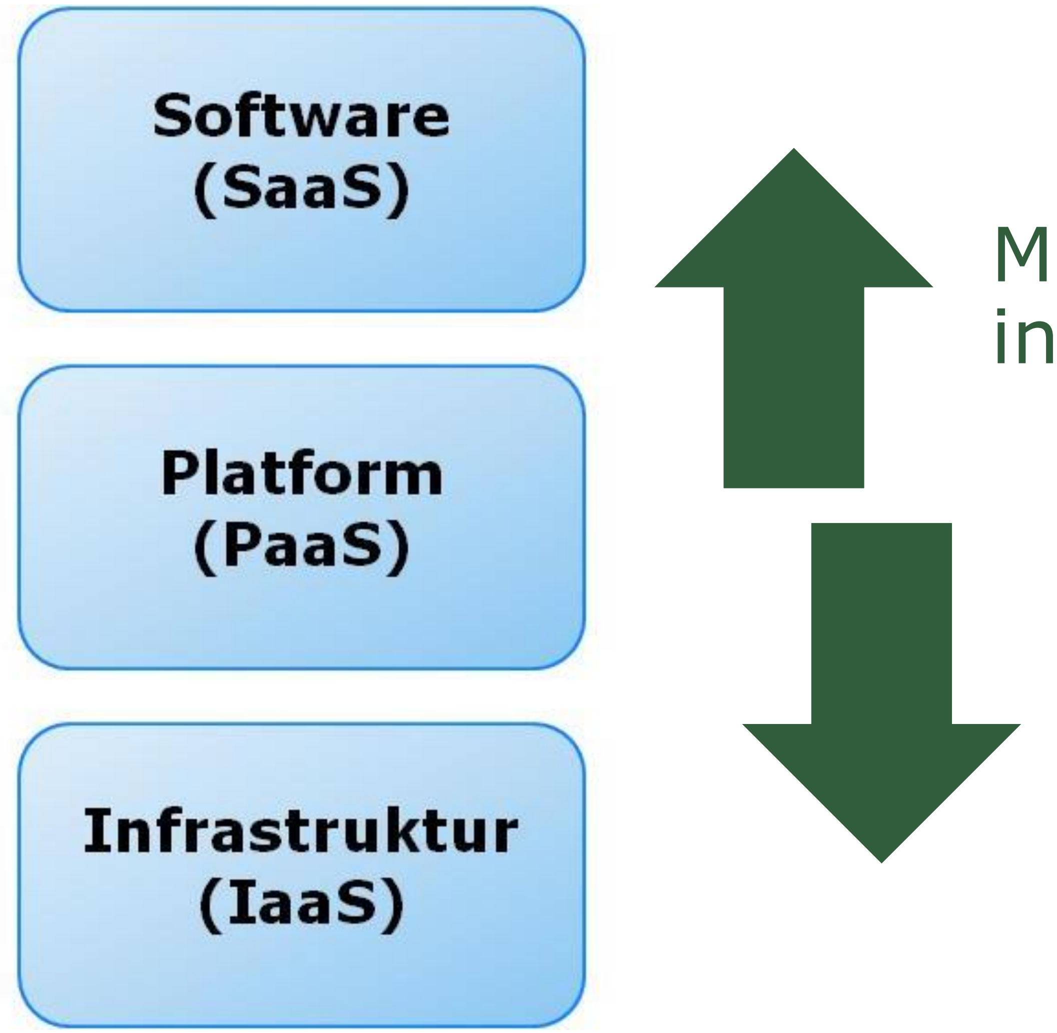
On-demand selvbetjening

Elasticitet

Betalingsmodellerne

IaaS: Ops without hardware
PaaS: Devs without Ops
SaaS: Business without Devs

De tre *aaS modeller



Software (SaaS)

Platform (PaaS)

Infrastruktur (IaaS)

”Fri for”

Mindre indflydelse

Mulighed for mere indflydelse

”Fri til”

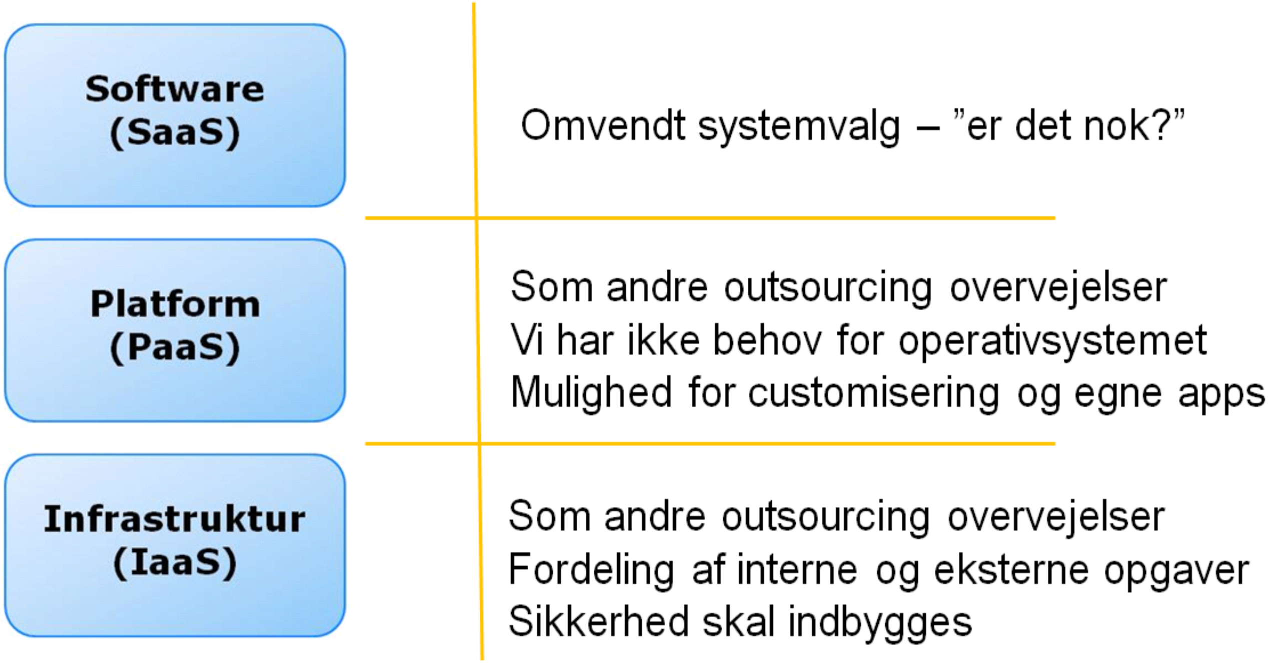
AWS vs Azure vs GoogleCloud vs Alibaba osv

Delt ansvar…

Løsning:	Eget ansvar:	Cloud-leverandørs ansvar:
Software (SaaS)	Konfiguration af log	Data Applikationer
Platform (PaaS)	Logs fra egne apps	System Management
Infrastruktur (IaaS)	Lokal overvågning Applikationslogs OS logs	Netværk Hardware, host Procedurer m.m. Fysisk sikring

Overvejelserne

De fleste overvejelser i forbindelse med outsourcing gælder også for cloudsourcing



Software (SaaS)

Omvendt systemvalg – ”er det nok?”

Platform (PaaS)

Som andre outsourcing overvejelser
Vi har ikke behov for operativsystemet
Mulighed for customisering og egne apps

Infrastruktur (IaaS)

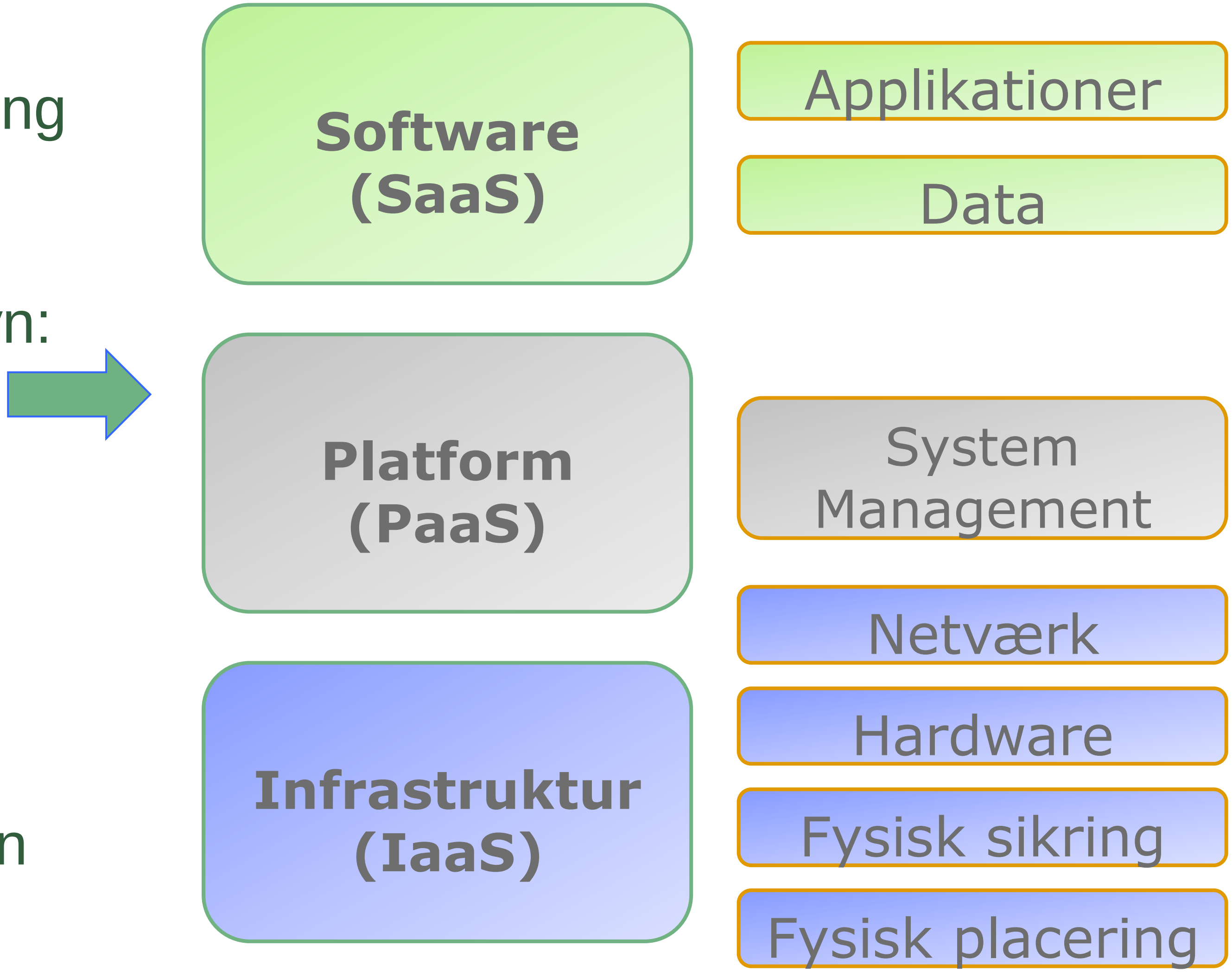
Som andre outsourcing overvejelser
Fordeling af interne og eksterne opgaver
Sikkerhed skal indbygges

Arbejdsgang - risikovurdering

Lovkrav:
Persondatalovgivning
Regnskabsloven

Compliance hensyn:
PCI
SOX
ISO 27001

Risiko vurdering
+
Data klassifikation



Software (SaaS)

Applikationer

Data

Platform (PaaS)

System Management

Infrastruktur (IaaS)

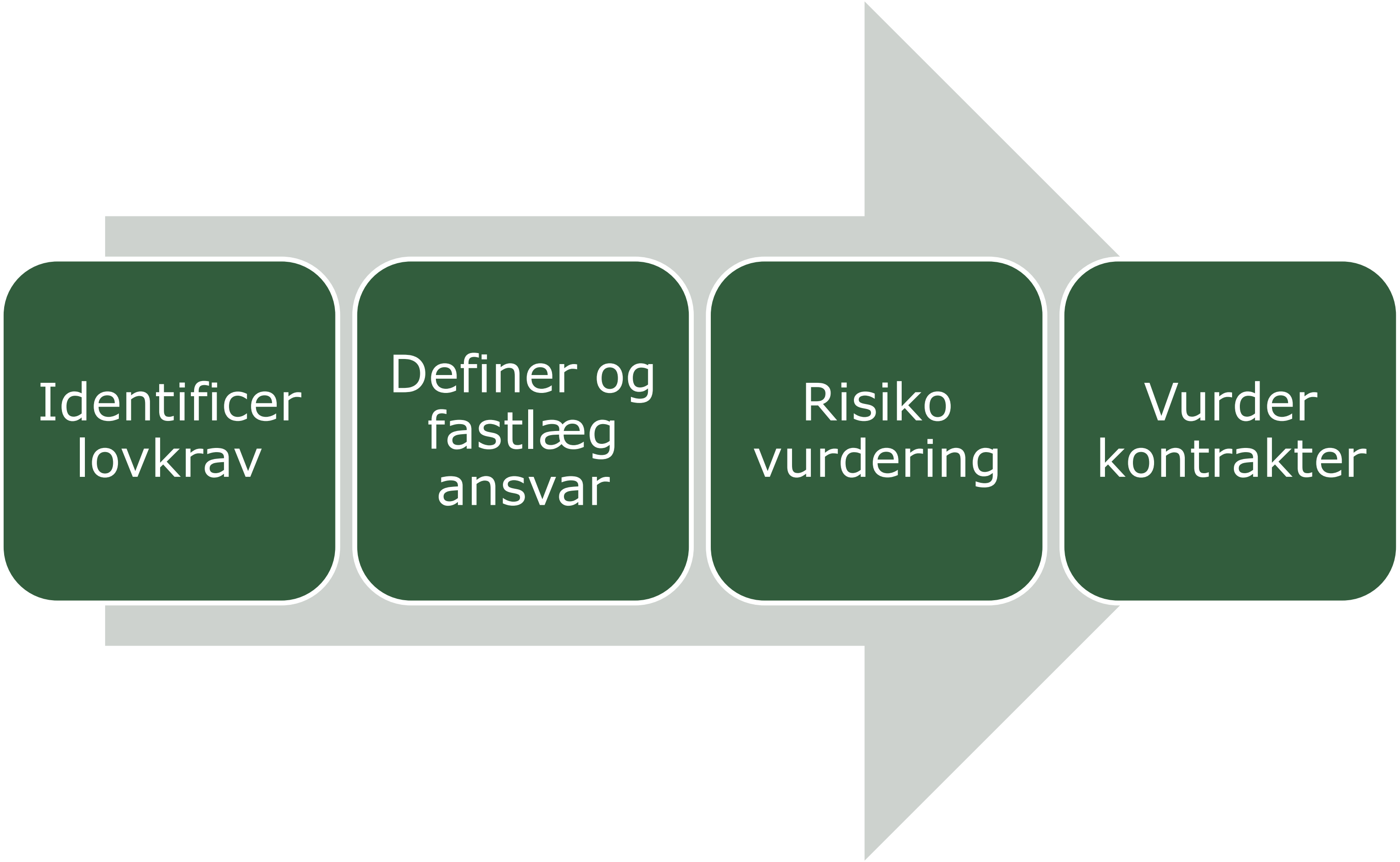
Netværk

Hardware

Fysisk sikring

Fysisk placering

Interne cloud krav



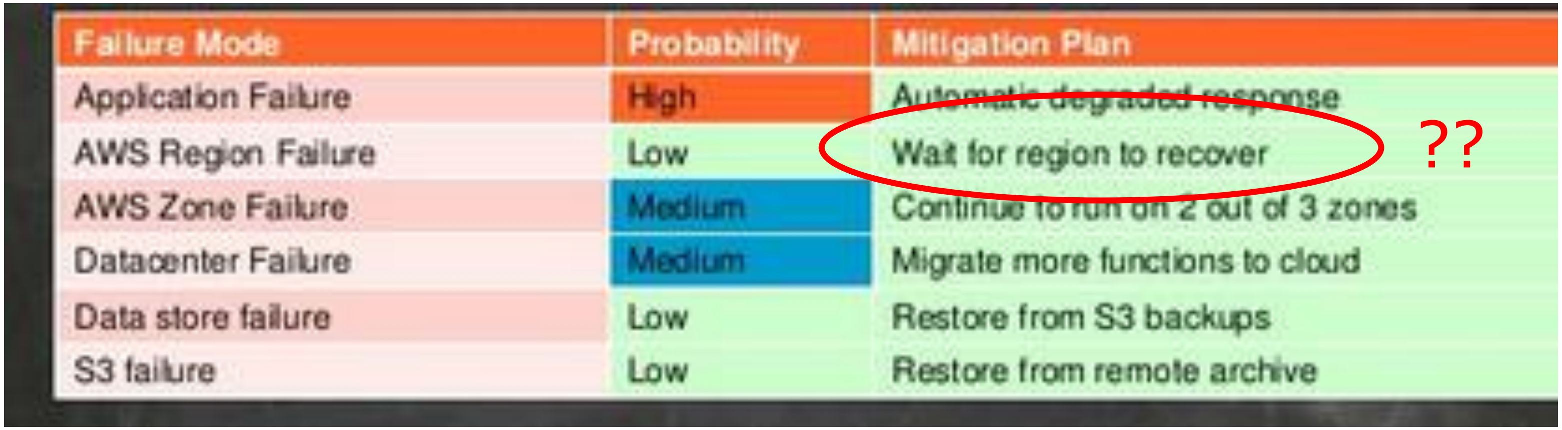
Identificer lovkrav

Definer og fastlæg ansvar

Risiko vurdering

Vurder kontrakter

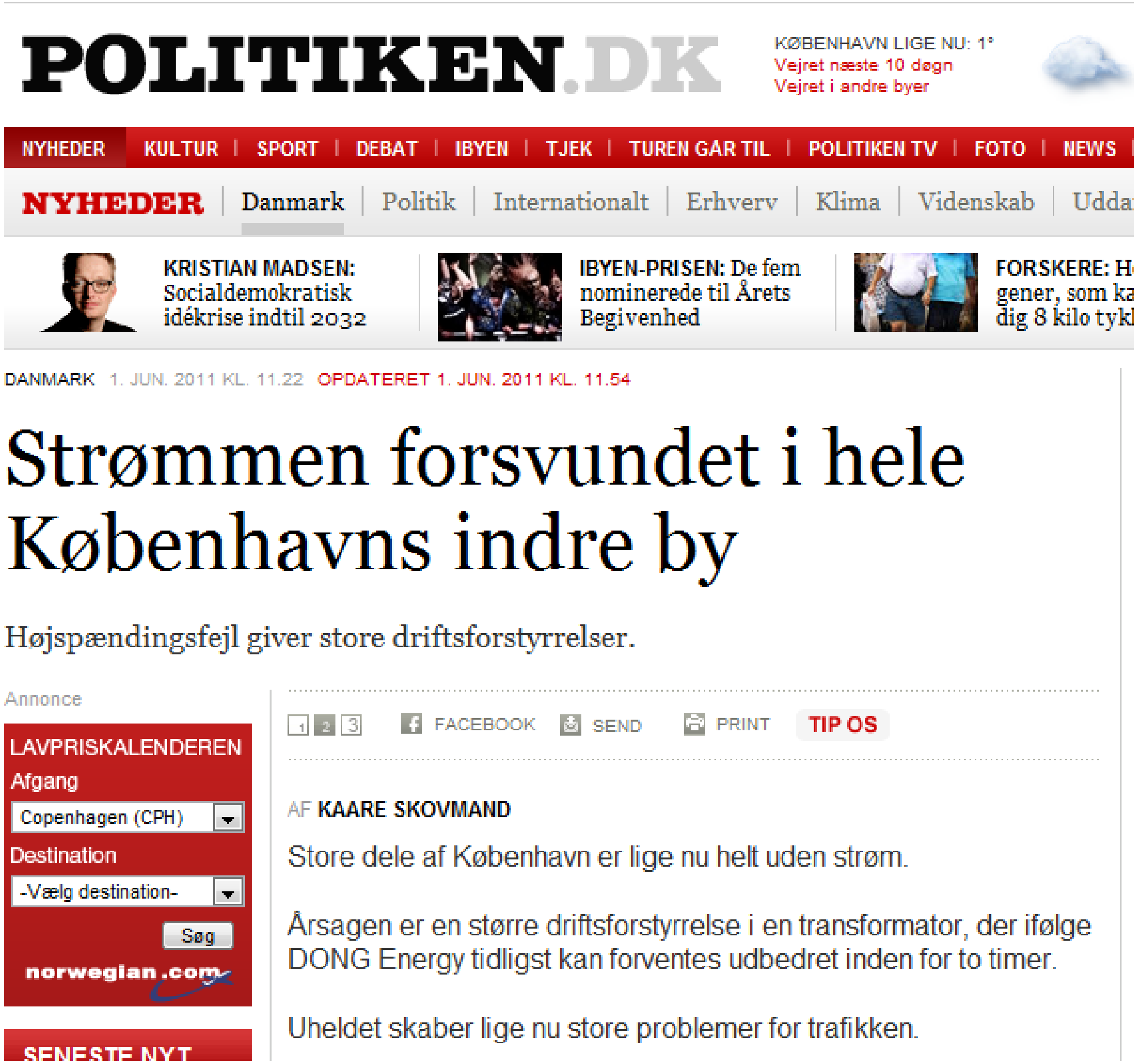
Cloud risikovurdering



Failure Mode	Probability	Mitigation Plan
Application Failure	High	Automatic degraded response
AWS Region Failure	Low	Wait for region to recover
AWS Zone Failure	Medium	Continue to run on 2 out of 3 zones
Datacenter Failure	Medium	Migrate more functions to cloud
Data store failure	Low	Restore from S3 backups
S3 failure	Low	Restore from remote archive

??

Men ikke meget anderledes



POLITIKEN.DK

KØBENHAVN LIGE NU: 1°
Vejret næste 10 døgn
Vejret i andre byer

NYHEDER | KULTUR | SPORT | DEBAT | IBYEN | TJEK | TUREN GÅR TIL | POLITIKEN TV | FOTO | NEWS
NYHEDER | Danmark | Politik | Internationalt | Erhverv | Klima | Videnskab | Udda

KRISTIAN MADSEN: Socialdemokratisk idékrise indtil 2032

IBYEN-PRISEN: De fem nominerede til Årets Begivenhed

FORSKERE: H gener, som ka dig 8 kilo tykl

DANMARK 1. JUN. 2011 KL. 11.22 OPDATERET 1. JUN. 2011 KL. 11.54

Strømmen forsvundet i hele Københavns indre by

Højspændingsfejl giver store driftsforstyrrelser.

AF KAARE SKOVMAND

Store dele af København er lige nu helt uden strøm.

Årsagen er en større driftsforstyrrelse i en transformator, der ifølge DONG Energy tidligst kan forventes udbedret inden for to timer.

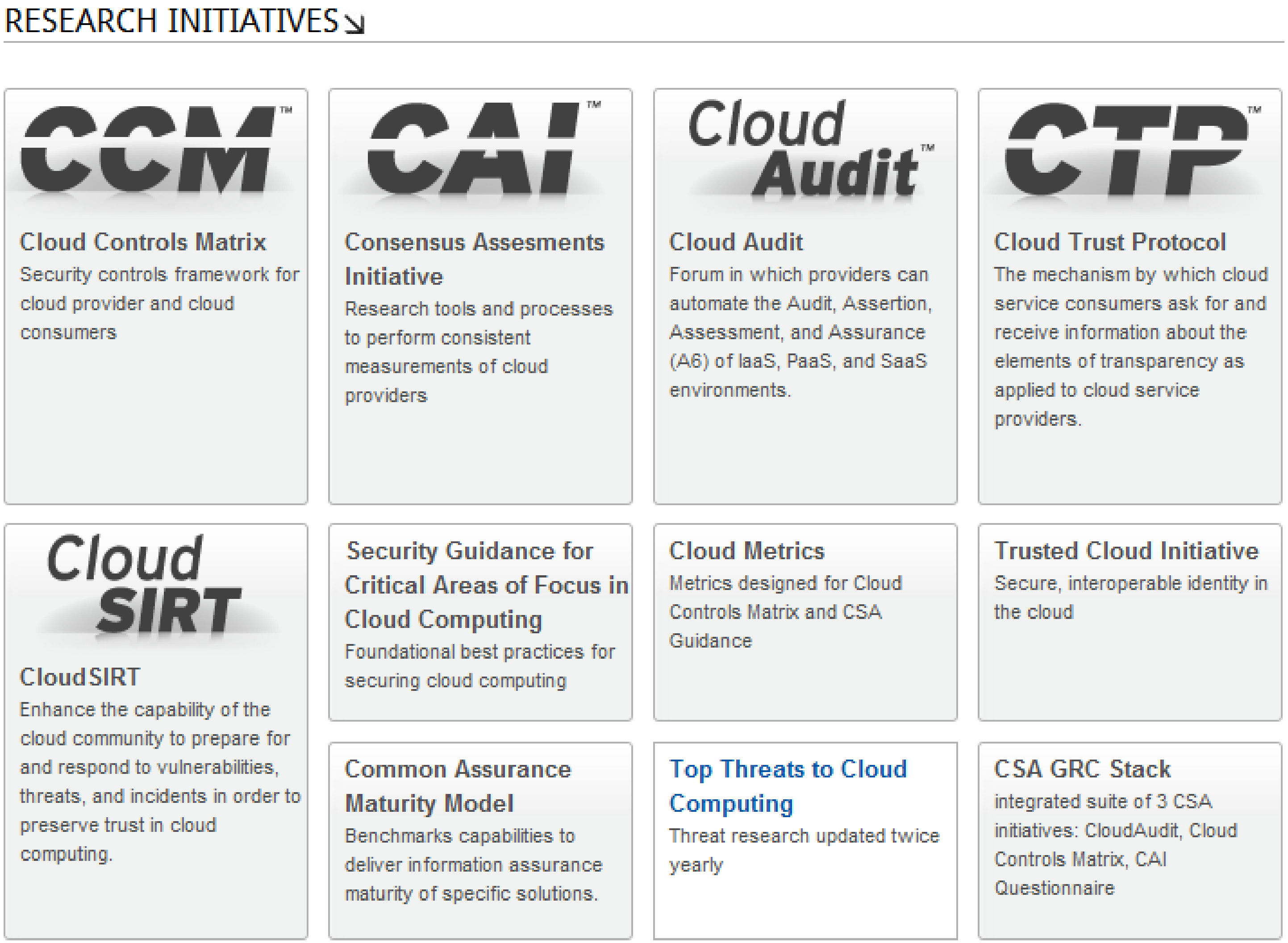
Uheldet skaber lige nu store problemer for trafikken.

Ikke magi

Det er ikke nødvendigt at starte forfra på cloud sikkerhedsarbejdet, mine sikkerhedskrav er (nok) ikke unikke

ISACA
Trust in, and value from, information systems
amazon web services™
Google Apps
talk
CSA cloud security alliance
Windows Azure
salesforce
SOFTWARE
enisa
European Network and Information Security Agency
iSF Information Security Forum
NIST
National Institute of Standards and Technology
U.S. Department of Commerce

cloudsecurityalliance.org



RESEARCH INITIATIVES ↘

Cloud Controls Matrix
Security controls framework for cloud provider and cloud consumers

Consensus Assesments Initiative
Research tools and processes to perform consistent measurements of cloud providers

Cloud Audit
Forum in which providers can automate the Audit, Assertion, Assessment, and Assurance (A6) of IaaS, PaaS, and SaaS environments.

Cloud Trust Protocol
The mechanism by which cloud service consumers ask for and receive information about the elements of transparency as applied to cloud service providers.

CloudSIRT
Enhance the capability of the cloud community to prepare for and respond to vulnerabilities, threats, and incidents in order to preserve trust in cloud computing.

Security Guidance for Critical Areas of Focus in Cloud Computing
Foundational best practices for securing cloud computing

Cloud Metrics
Metrics designed for Cloud Controls Matrix and CSA Guidance

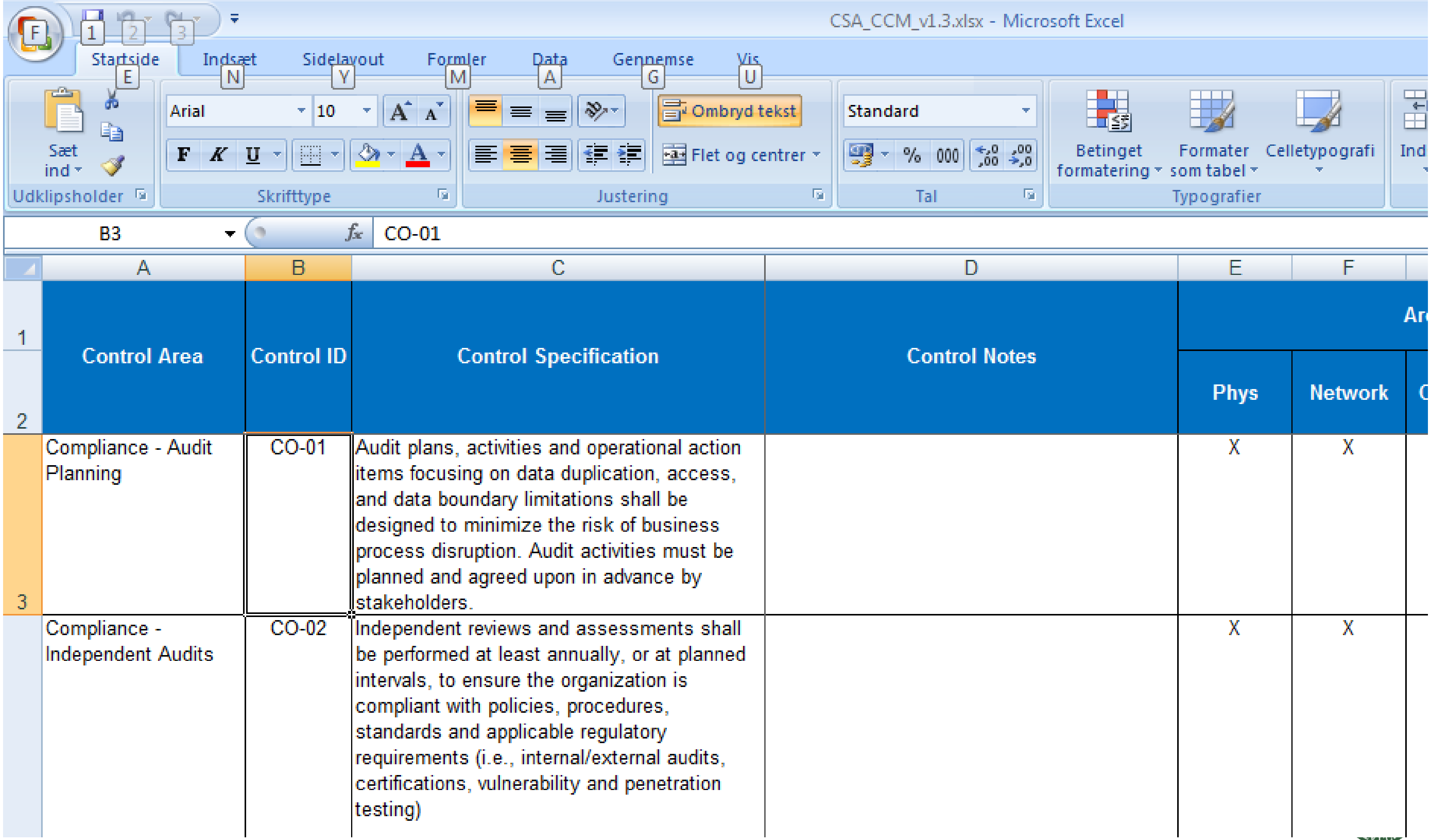
Trusted Cloud Initiative
Secure, interoperable identity in the cloud

Common Assurance Maturity Model
Benchmarks capabilities to deliver information assurance maturity of specific solutions.

Top Threats to Cloud Computing
Threat research updated twice yearly

CSA GRC Stack
integrated suite of 3 CSA initiatives: CloudAudit, Cloud Controls Matrix, CAI Questionnaire

Cloud Audit – Cloud Controls Matrix

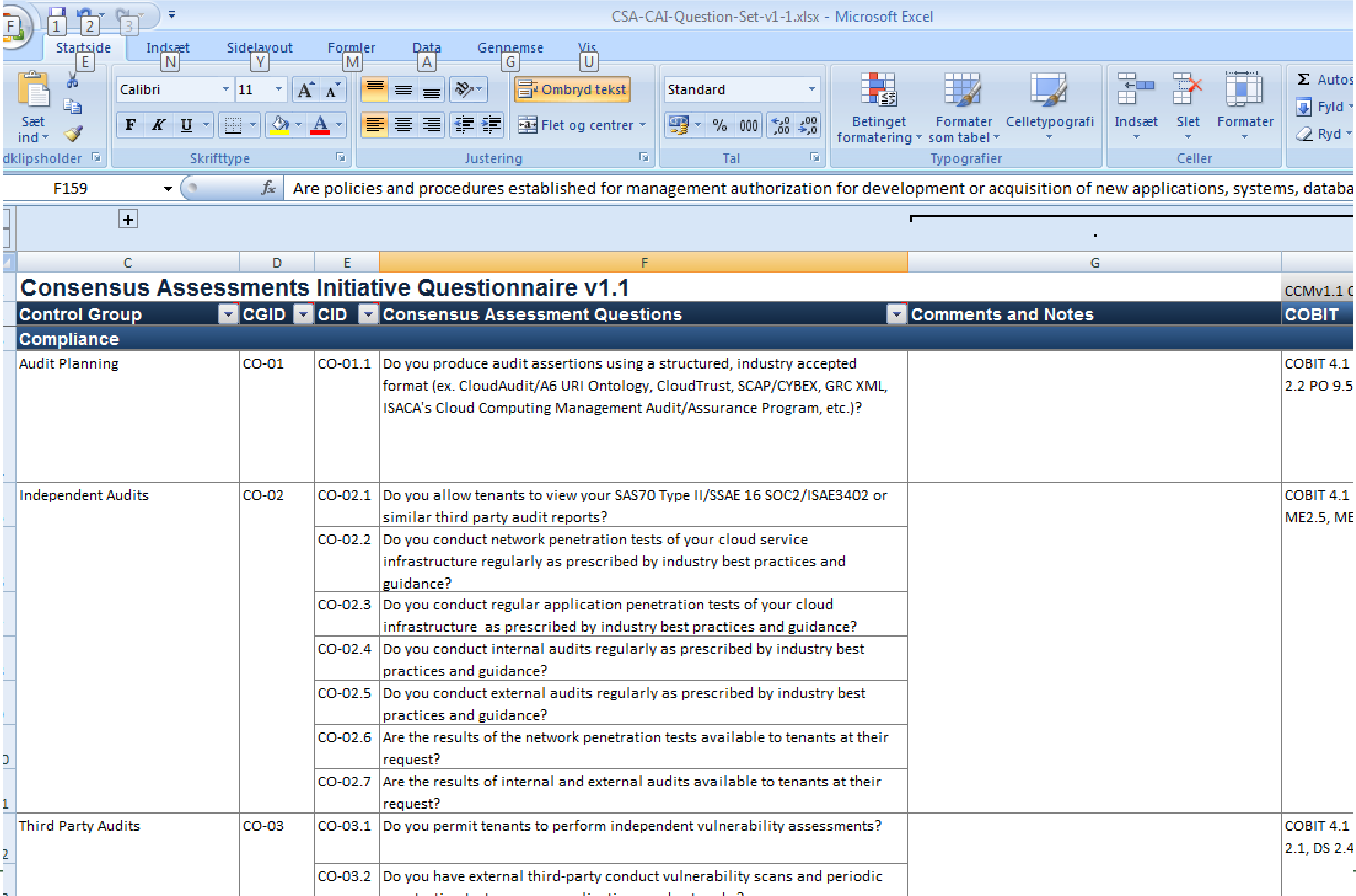


CSA_CCM_v1.3.xlsx - Microsoft Excel

B3 | *fx* | CO-01

Control Area	Control ID	Control Specification	Control Notes		Ar	
				Phys	Network	C
Compliance - Audit Planning	CO-01	Audit plans, activities and operational action items focusing on data duplication, access, and data boundary limitations shall be designed to minimize the risk of business process disruption. Audit activities must be planned and agreed upon in advance by stakeholders.		X	X	
Compliance - Independent Audits	CO-02	Independent reviews and assessments shall be performed at least annually, or at planned intervals, to ensure the organization is compliant with policies, procedures, standards and applicable regulatory requirements (i.e., internal/external audits, certifications, vulnerability and penetration testing)		X	X	

Cloud Audit – Consensus Assessment Initiative



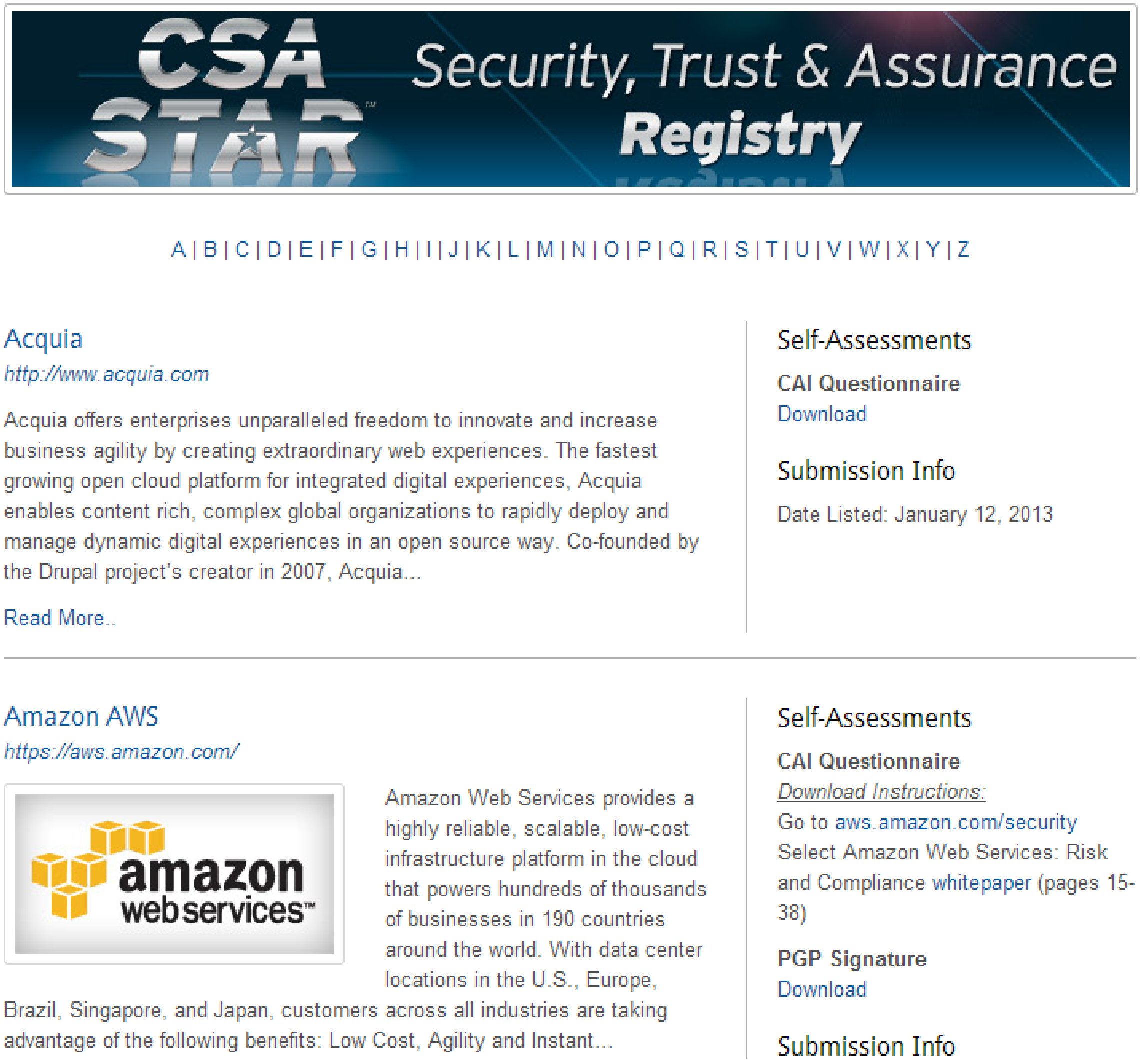
F159 | Are policies and procedures established for management authorization for development or acquisition of new applications, systems, databa

Consensus Assessments Initiative Questionnaire v1.1

CCMv1.1 C

Control Group	CGID	CID	Consensus Assessment Questions	Comments and Notes	COBIT
Compliance					
Audit Planning	CO-01	CO-01.1	Do you produce audit assertions using a structured, industry accepted format (ex. CloudAudit/A6 URI Ontology, CloudTrust, SCAP/CYBEX, GRC XML, ISACA's Cloud Computing Management Audit/Assurance Program, etc.)?		COBIT 4.1 2.2 PO 9.5
Independent Audits	CO-02	CO-02.1	Do you allow tenants to view your SAS70 Type II/SSAE 16 SOC2/ISAE3402 or similar third party audit reports?		COBIT 4.1 ME2.5, ME
		CO-02.2	Do you conduct network penetration tests of your cloud service infrastructure regularly as prescribed by industry best practices and guidance?		
		CO-02.3	Do you conduct regular application penetration tests of your cloud infrastructure as prescribed by industry best practices and guidance?		
		CO-02.4	Do you conduct internal audits regularly as prescribed by industry best practices and guidance?		
		CO-02.5	Do you conduct external audits regularly as prescribed by industry best practices and guidance?		
		CO-02.6	Are the results of the network penetration tests available to tenants at their request?		
		CO-02.7	Are the results of internal and external audits available to tenants at their request?		
Third Party Audits	CO-03	CO-03.1	Do you permit tenants to perform independent vulnerability assessments?		COBIT 4.1 2.1, DS 2.4
		CO-03.2	Do you have external third-party conduct vulnerability scans and periodic		

cloudsecurityalliance.org/star/registry



A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z

Acquia

http://www.acquia.com

Acquia offers enterprises unparalleled freedom to innovate and increase business agility by creating extraordinary web experiences. The fastest growing open cloud platform for integrated digital experiences, Acquia enables content rich, complex global organizations to rapidly deploy and manage dynamic digital experiences in an open source way. Co-founded by the Drupal project's creator in 2007, Acquia...

Read More..

Self-Assessments

CAI Questionnaire

Download

Submission Info

Date Listed: January 12, 2013

Amazon AWS

https://aws.amazon.com/

Amazon Web Services provides a highly reliable, scalable, low-cost infrastructure platform in the cloud that powers hundreds of thousands of businesses in 190 countries around the world. With data center locations in the U.S., Europe, Brazil, Singapore, and Japan, customers across all industries are taking advantage of the following benefits: Low Cost, Agility and Instant...

Self-Assessments

CAI Questionnaire

Download Instructions:

Go to aws.amazon.com/security

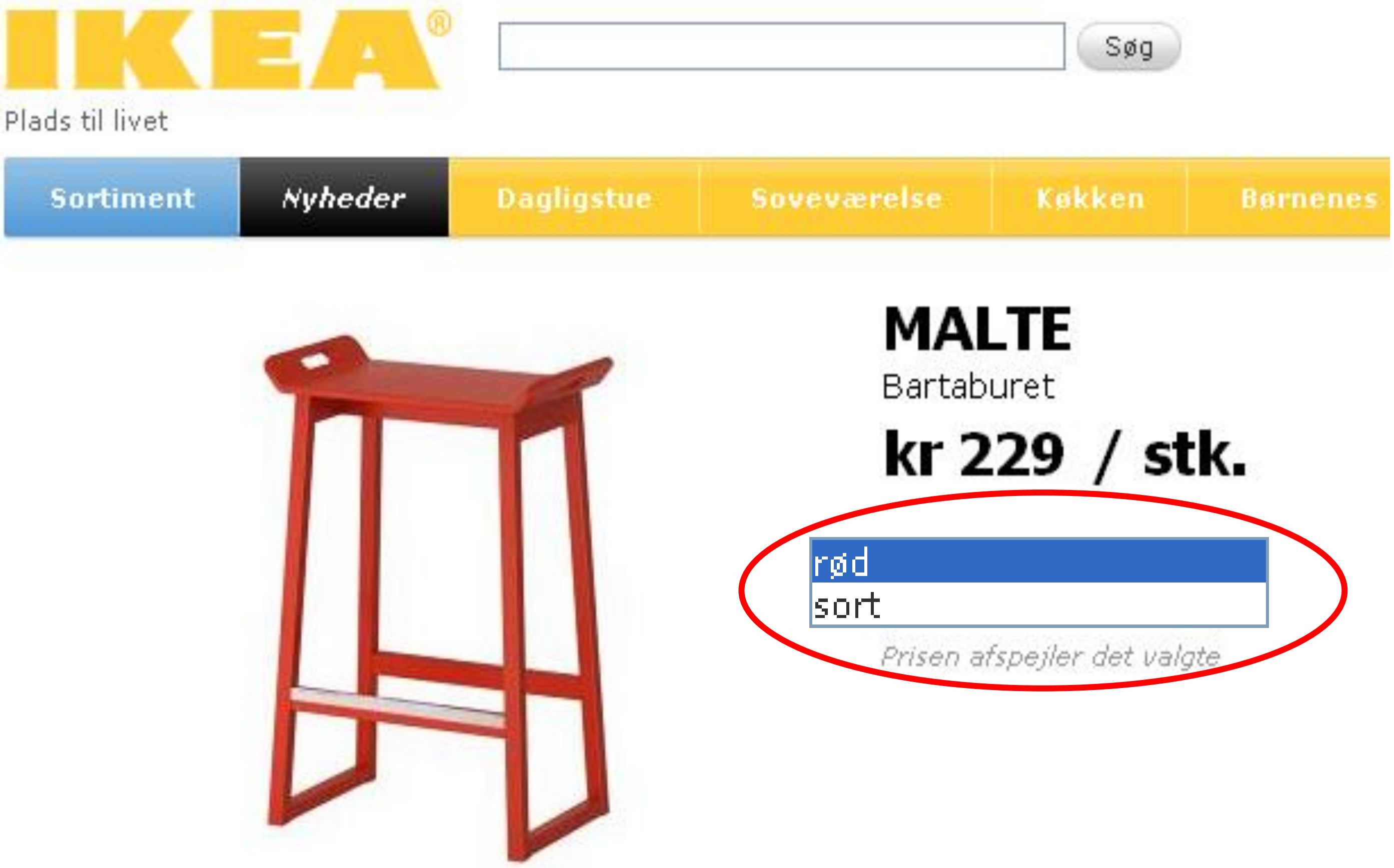
Select Amazon Web Services: Risk and Compliance whitepaper (pages 15-38)

PGP Signature

Download

Submission Info

Kan jeg få den i grøn?



MALTE

Bartaburet

kr 229 / stk.

rød
sort

Prisen afspejler det valgte

”Cloud”, cloud eller CLOUD

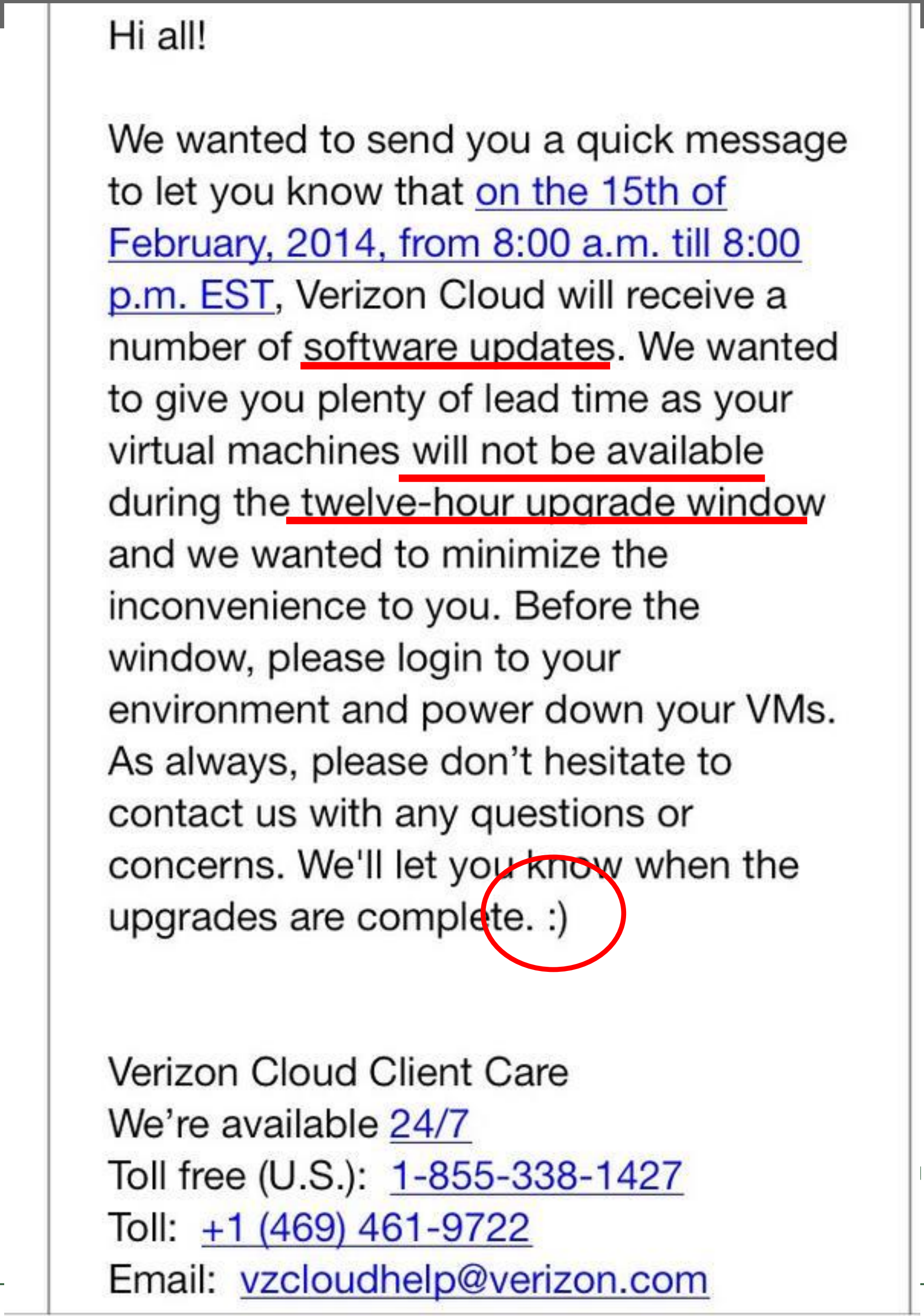
eller

”Cloud” og cloud – traditionel outsourcing eller cloudsourcing

Standardisering

Ingen IT i 12 timer :)

Clouds er forskellige



Hi all!

We wanted to send you a quick message to let you know that on the 15th of February, 2014, from 8:00 a.m. till 8:00 p.m. EST, Verizon Cloud will receive a number of software updates. We wanted to give you plenty of lead time as your virtual machines will not be available during the twelve-hour upgrade window and we wanted to minimize the inconvenience to you. Before the window, please login to your environment and power down your VMs. As always, please don't hesitate to contact us with any questions or concerns. We'll let you know when the upgrades are complete. :)

Verizon Cloud Client Care
We're available 24/7
Toll free (U.S.): 1-855-338-1427
Toll: +1 (469) 461-9722
Email: vzcloudhelp@verizon.com

Sikkerhed i skyen

Alle de kendte sikkerhedsudfordinger findes i skyen

Din kontrakt – og lovgivningen

Det er **DIT** ansvar at vælge en leverandør, der leverer den fornødne grad af teknisk sikkerhed og forsvarlige procedurer, og det er **DIT** ansvar at kontrollere overholdelsen af det aftalte.

Data i EU

Brugen af kryptering

Leverandørens muligheder for adgang til din data

Registreredes rettigheder

CLOUD Act, Patriot Act, FISAA…

Cloud sikkerhed >< traditionel it-sikkerhed

- ”Design for failures” – forvent service issues
- Opdater og udrul nye instanser, ikke de kørende
- Paranoid arkitektur: opdel services
- Kryptering, data at rest

Cloud sikkerhed og traditional sikkerhed

To-faktor adgang

Brug af begrænsede konti fra starten, også i cloud

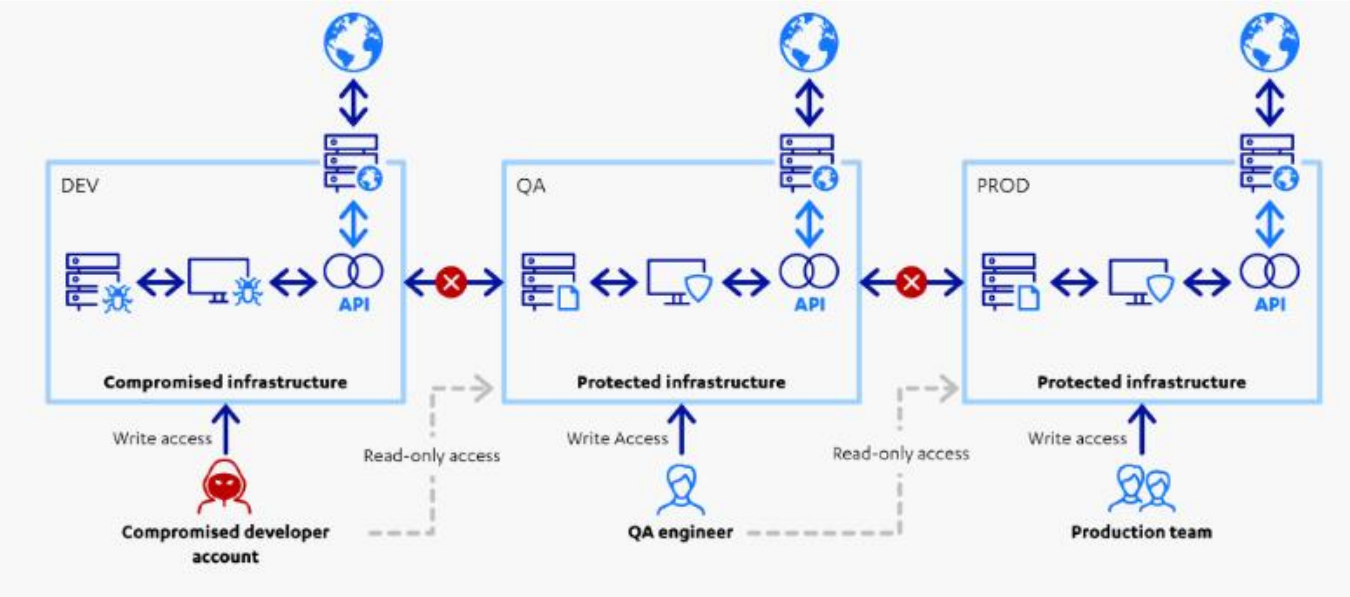
Brug forskellige sikkerhedsgrupper, adskilte admingrupper og sikkerhedsgrupper

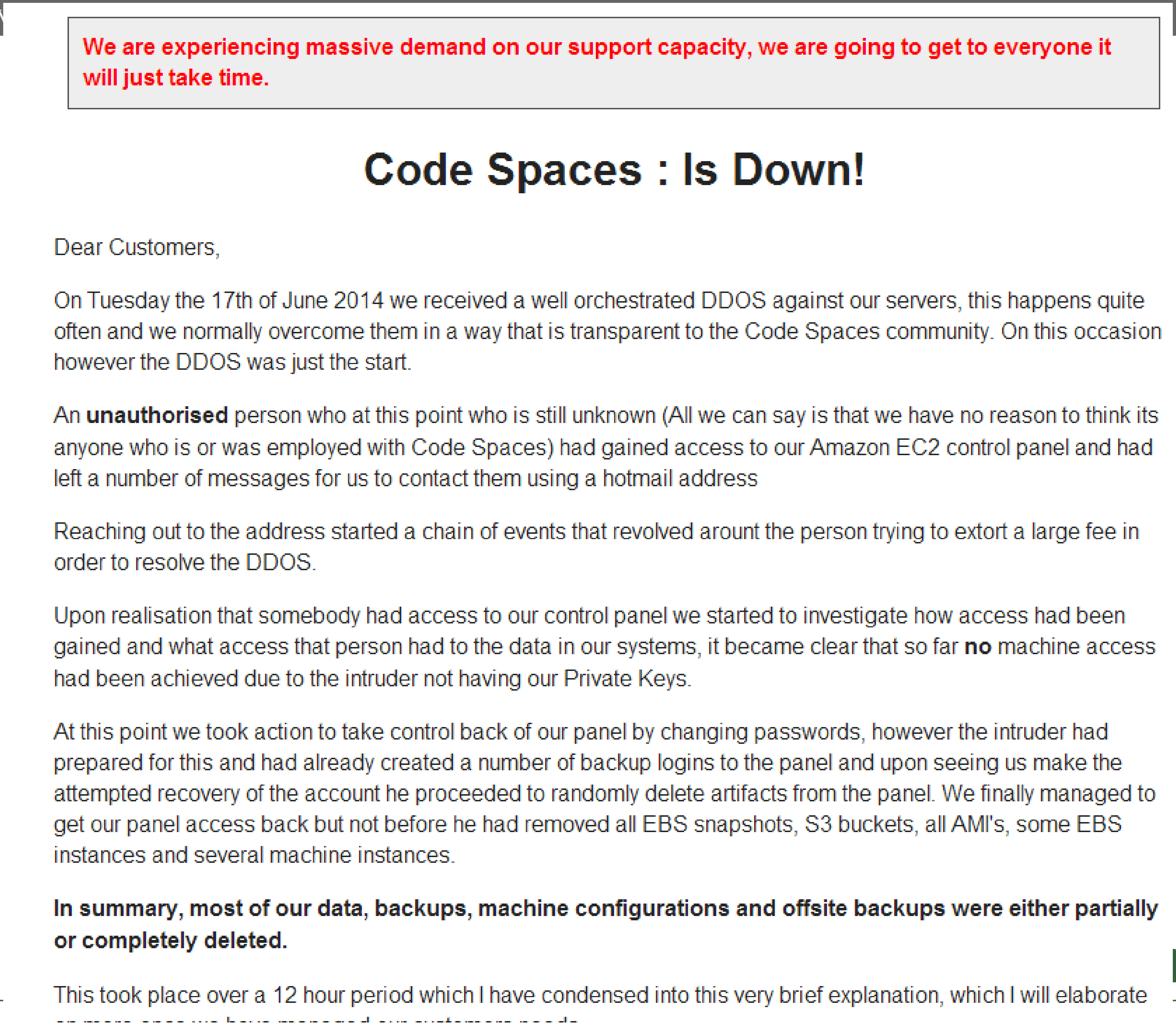
Hvad sker der når cloud-løsningen fejler

Single Point of Failures og afhængigheder ved kombineret infrastruktur

- Availability
- Laveste fællesnævner
- Delvis tilgængelighed

Reducer ”blast radius”





We are experiencing massive demand on our support capacity, we are going to get to everyone it will just take time.

Code Spaces : Is Down!

Dear Customers,

On Tuesday the 17th of June 2014 we received a well orchestrated DDOS against our servers, this happens quite often and we normally overcome them in a way that is transparent to the Code Spaces community. On this occasion however the DDOS was just the start.

An **unauthorised** person who at this point who is still unknown (All we can say is that we have no reason to think its anyone who is or was employed with Code Spaces) had gained access to our Amazon EC2 control panel and had left a number of messages for us to contact them using a hotmail address

Reaching out to the address started a chain of events that revolved arount the person trying to extort a large fee in order to resolve the DDOS.

Upon realisation that somebody had access to our control panel we started to investigate how access had been gained and what access that person had to the data in our systems, it became clear that so far **no** machine access had been achieved due to the intruder not having our Private Keys.

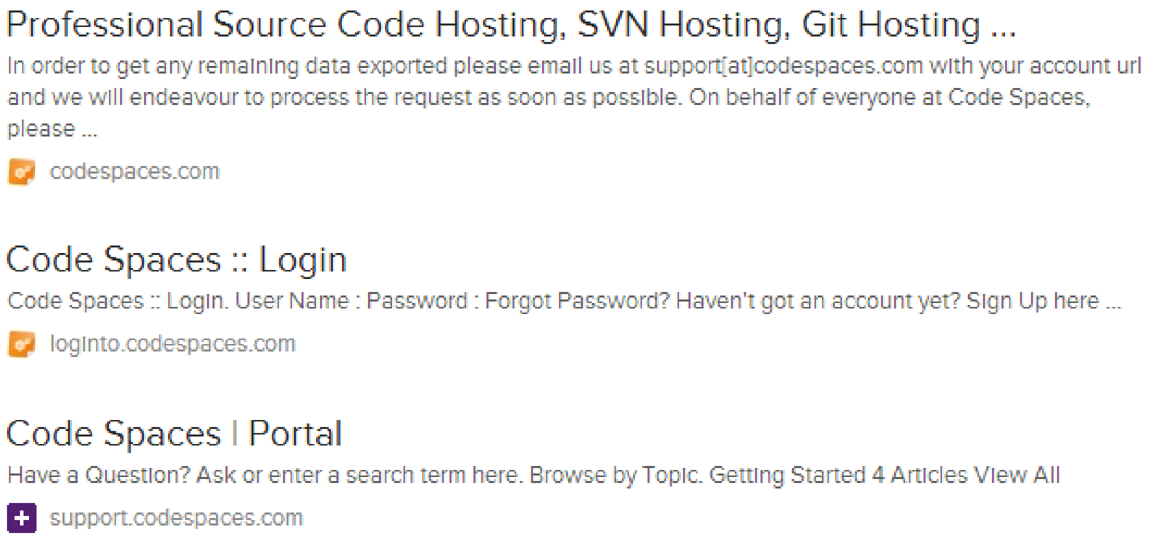
At this point we took action to take control back of our panel by changing passwords, however the intruder had prepared for this and had already created a number of backup logins to the panel and upon seeing us make the attempted recovery of the account he proceeded to randomly delete artifacts from the panel. We finally managed to get our panel access back but not before he had removed all EBS snapshots, S3 buckets, all AMI's, some EBS instances and several machine instances.

In summary, most of our data, backups, machine configurations and offsite backups were either partially or completely deleted.

This took place over a 12 hour period which I have condensed into this very brief explanation, which I will elaborate on more once we have managed our customers needs.

Codespaces.com

The attacker deleted
“all machine [VMs], all EBS vols containing database files, all snapshots & backups, and all S3 data”.



Professional Source Code Hosting, SVN Hosting, Git Hosting ...
In order to get any remaining data exported please email us at support[at]codespaces.com with your account url and we will endeavour to process the request as soon as possible. On behalf of everyone at Code Spaces, please ...
codespaces.com

Code Spaces :: Login
Code Spaces :: Login. User Name : Password : Forgot Password? Haven't got an account yet? Sign Up here ...
loginto.codespaces.com

Code Spaces | Portal
Have a Question? Ask or enter a search term here. Browse by Topic. Getting Started 4 Articles View All
support.codespaces.com

Codespaces.com – some lessons

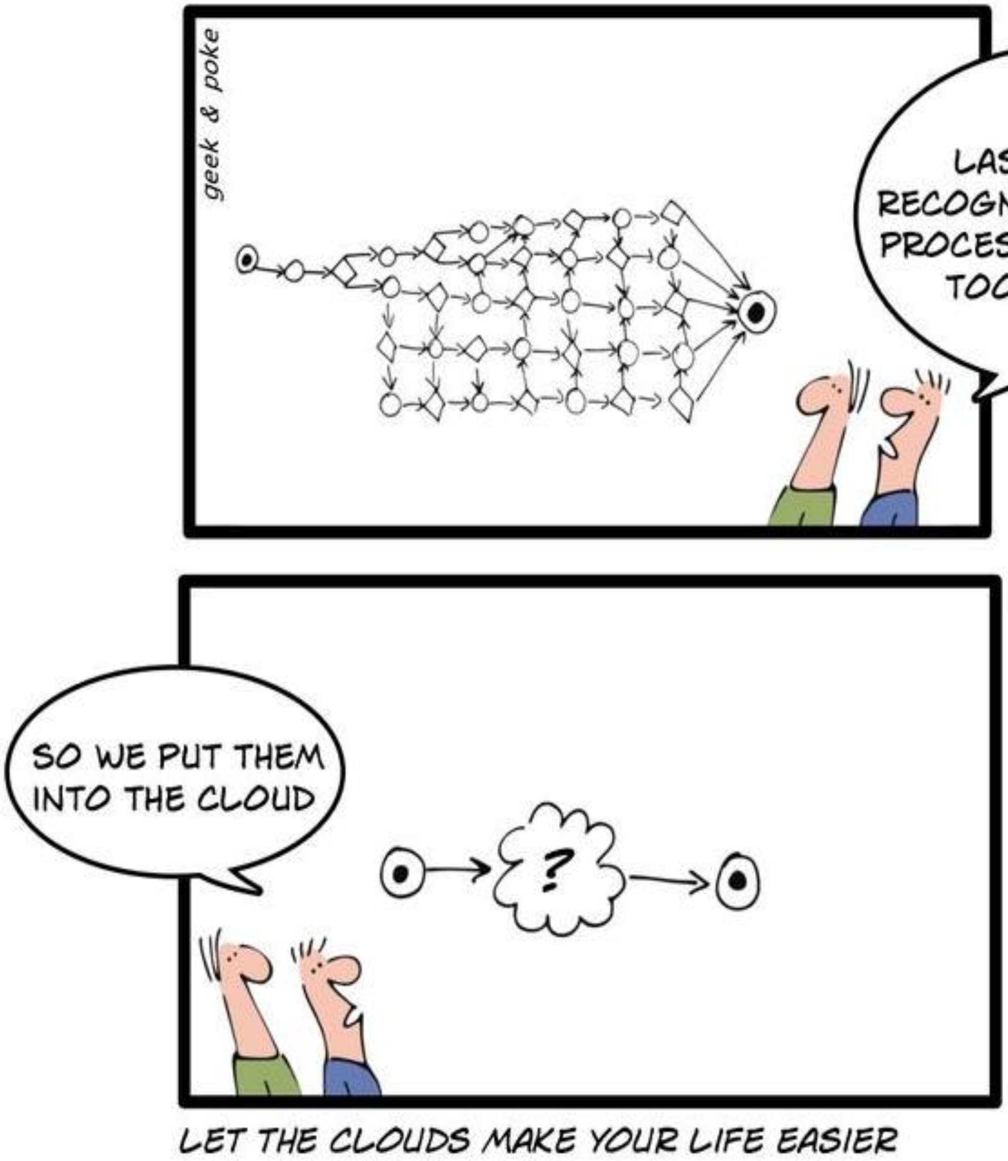
- Avoid using the master credential, use the Identity Management console
- Use Two Factor Authentication
- Segment backup access from the rest of the infrastructure. For instance backups could be archived into a different AWS account without delete access.

Du kan IKKE gøre skyen sikker

Men – med mindre du arbejder for en cloud-leverandør – skal du heller ikke.

Du skal kunne sikre dine **data** og dine **applikationer**

Let the clouds make your life easier


geek & poke
LAST YEAR WE RECOGNIZED THAT OUR PROCESSES WERE FAR TOO COMPLEX
SO WE PUT THEM INTO THE CLOUD
?
LET THE CLOUDS MAKE YOUR LIFE EASIER

Cloud Computing er helt normale it-systemer, der bruger strøm.

It-systemer fejler en gang imellem, men de kan vurderes.

Virtualisering og containers

Hvad er virtualisering ?

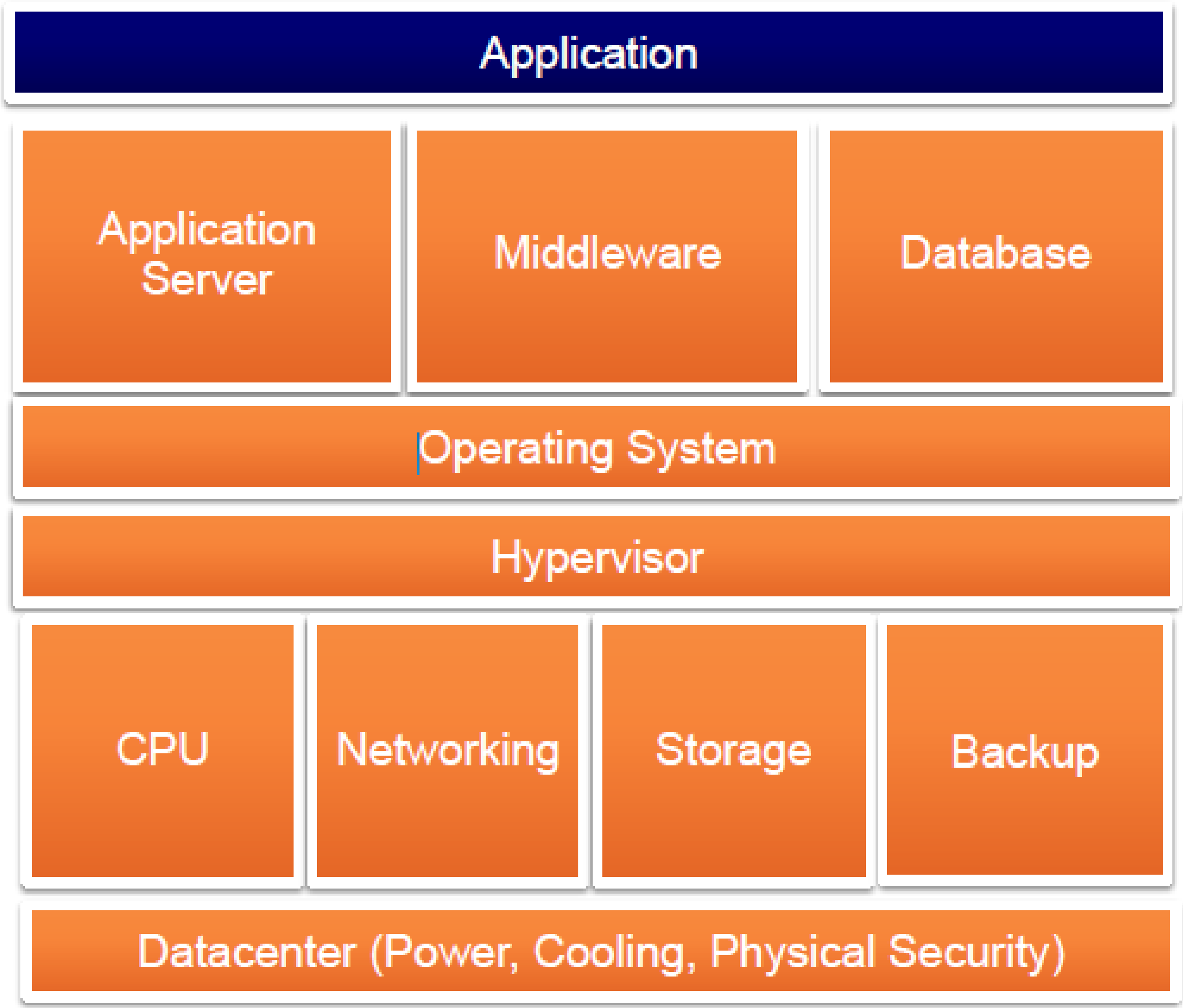
Hvad er virtualisering?

At få én fysisk enhed til at opføre sig som flere uafhængige enheder

Partitionere én fysisk server til flere ”virtuelle” servere, hvor hver ser ud til at køre som en dedikeret fysisk maskine. Hver server kan bootes uafhængigt af de andre.

Gæste operativ systemer/servere/storage

Hvad er virtualisering?



Application

Application Server

Middleware

Database

Operating System

Hypervisor

CPU

Networking

Storage

Backup

Datacenter (Power, Cooling, Physical Security)

Gæst

Vært

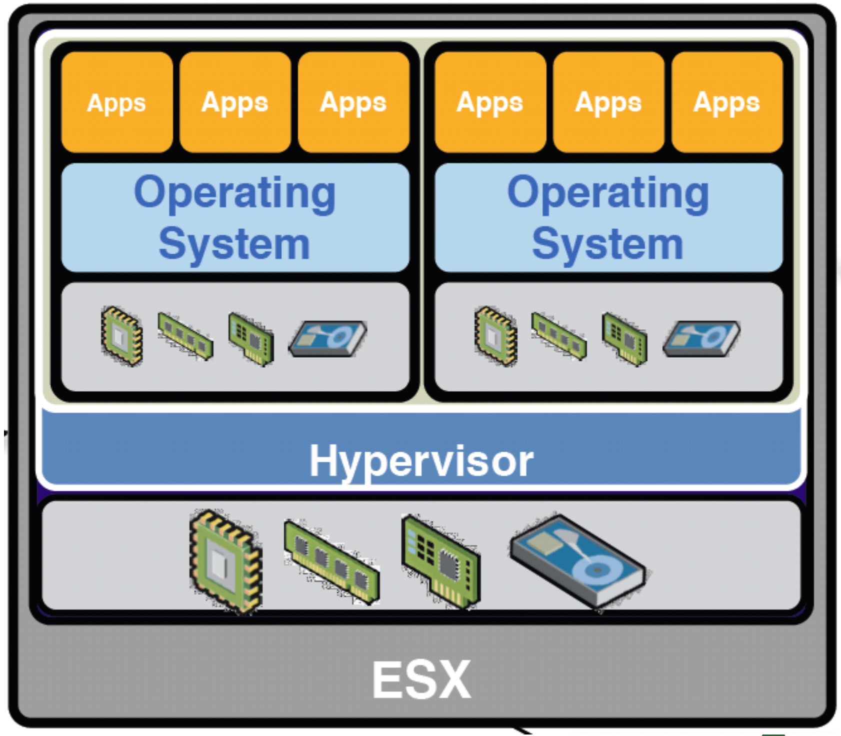
(Bare-metal
eller
Hosted)

Trusler imod virtualisering

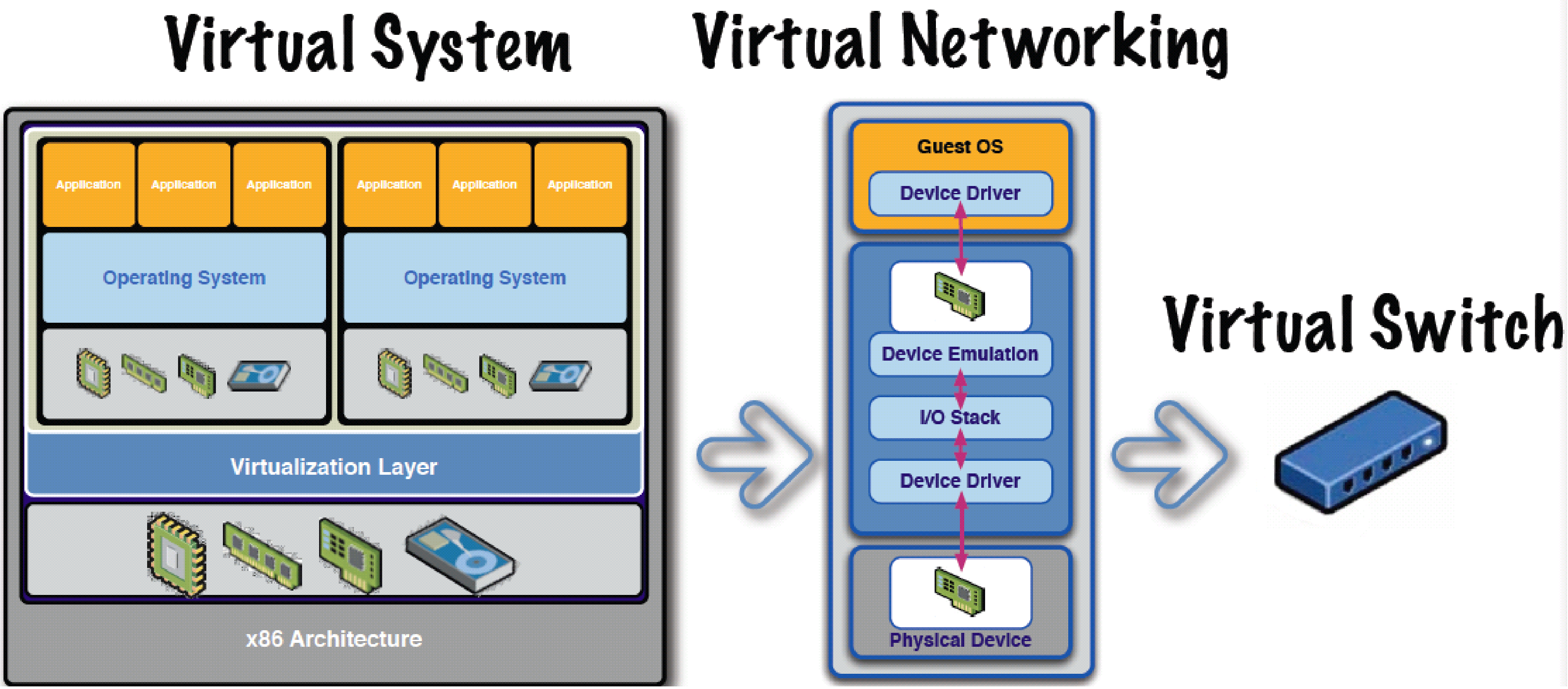
1. Guest to Self
2. Guest to Guest
3. Guest to Host/VMM/HW
4. External to Host/VMM/HW
5. External to Guest
6. Host/VMM to All…
7. Hardware to VMM

Administrationslaget…

Risikovurderingen



Trusler imod virtualisering



Trusler imod virtualisering

Sikkerhedslag kan flyttes til virtuliseringssoftwaren, f.eks. virusscanninger i hypervisoren

Udsætter hypervisor for potentielle risks

OS eller application layer: Host firewalls, AV, logning / log overvågning

Men - det påvirker performance og koster licenser, routning svært og beskytter ikke imod angreb inde fra de virtuelle miljøer

Virtualisering

Sikkerhedsproblemer opstår pga. fejlkonfiguration og dårligt design eller forkert implementering

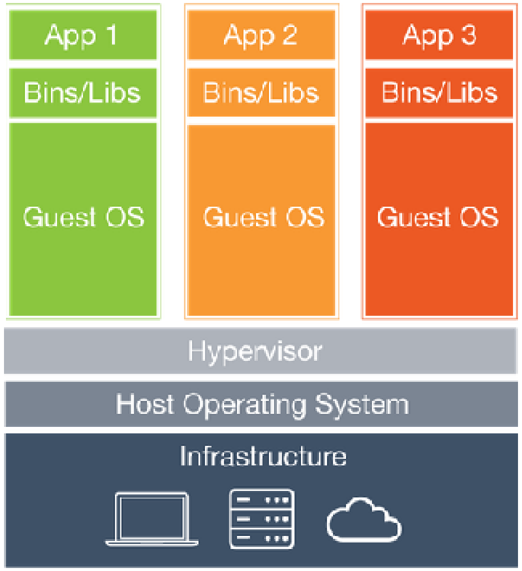
Alle leverandører har hærdningsvejledninger og best practice dokumentation

No free lunch

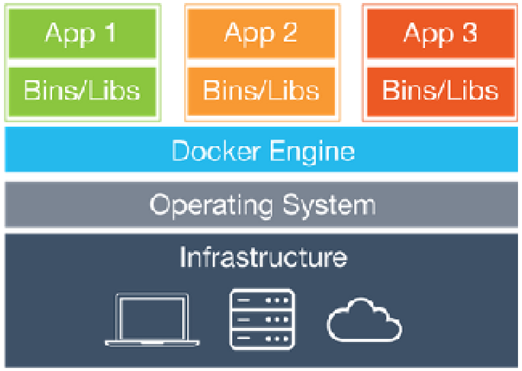
Containers and micro instances

Security isolation and application containment while improving resource efficiency over full virtual machines.

Linux containers provide segmentation via kernel namespaces, resource control via cgroups and are often secured through reduced root capabilities, Mandatory Access Control and user namespaces.



Virtual Machines



Containers

Containers and micro instances

Containers collapses the security perimeter

No layer 3 security, app sec takes over

Is the code running inside the container safe?

What has the container access to?

Who can it communicate with?

Where in the world is it running physically?

How is the container deployment and management ?

Pause

Co-worker: What do you mean be “hard coded” in the program?

Me:

Serverless og Legoklodser

Serverless…?

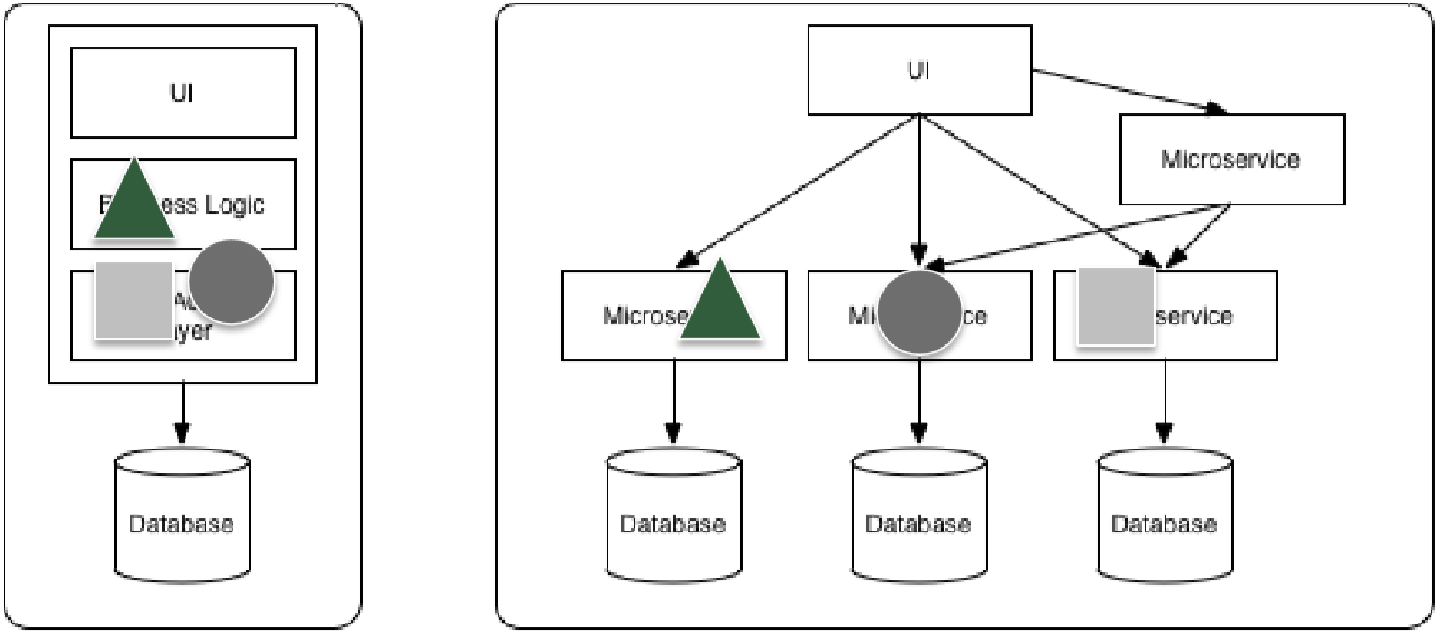
Ingen servere – ligesom der ikke er et køkken når du køber fastfood

Serverless…?

Eller – ligesom “Wireless” ikke har nogle kabler
(for dig, men der er mange, mange kabler bagved)

Ingen servere – ligesom der ikke er et køkken
når du køber fastfood

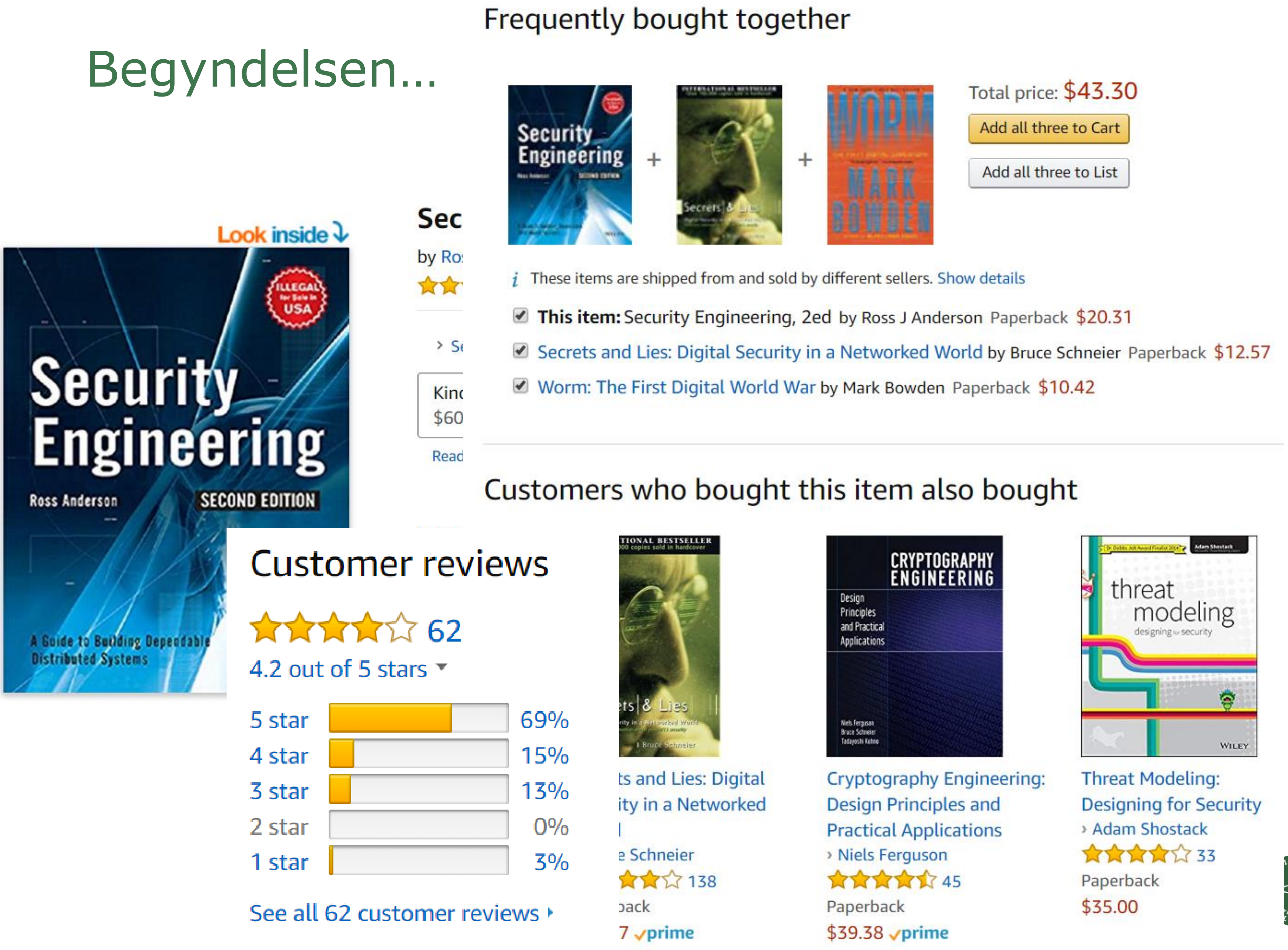
Begyndelsen – microservices…



Monolithic architecture

Microservices architecture

Begyndelsen…



Frequently bought together

Total price: $43.30

Add all three to Cart

Add all three to List

i These items are shipped from and sold by different sellers. Show details

☑ **This item:** Security Engineering, 2ed by Ross J Anderson Paperback $20.31

☑ Secrets and Lies: Digital Security in a Networked World by Bruce Schneier Paperback $12.57

☑ Worm: The First Digital World War by Mark Bowden Paperback $10.42

Customers who bought this item also bought

Customer reviews

★★★★☆ 62

4.2 out of 5 stars

5 star	69%
4 star	15%
3 star	13%
2 star	0%
1 star	3%

See all 62 customer reviews

ts and Lies: Digital ity in a Networked l
e Schneier
★★★☆☆ 138
pack
7 ✓prime

Cryptography Engineering: Design Principles and Practical Applications
› Niels Ferguson
★★★★☆ 45
Paperback
$39.38 ✓prime

Threat Modeling: Designing for Security
› Adam Shostack
★★★★☆ 33
Paperback
$35.00

Serverless...?

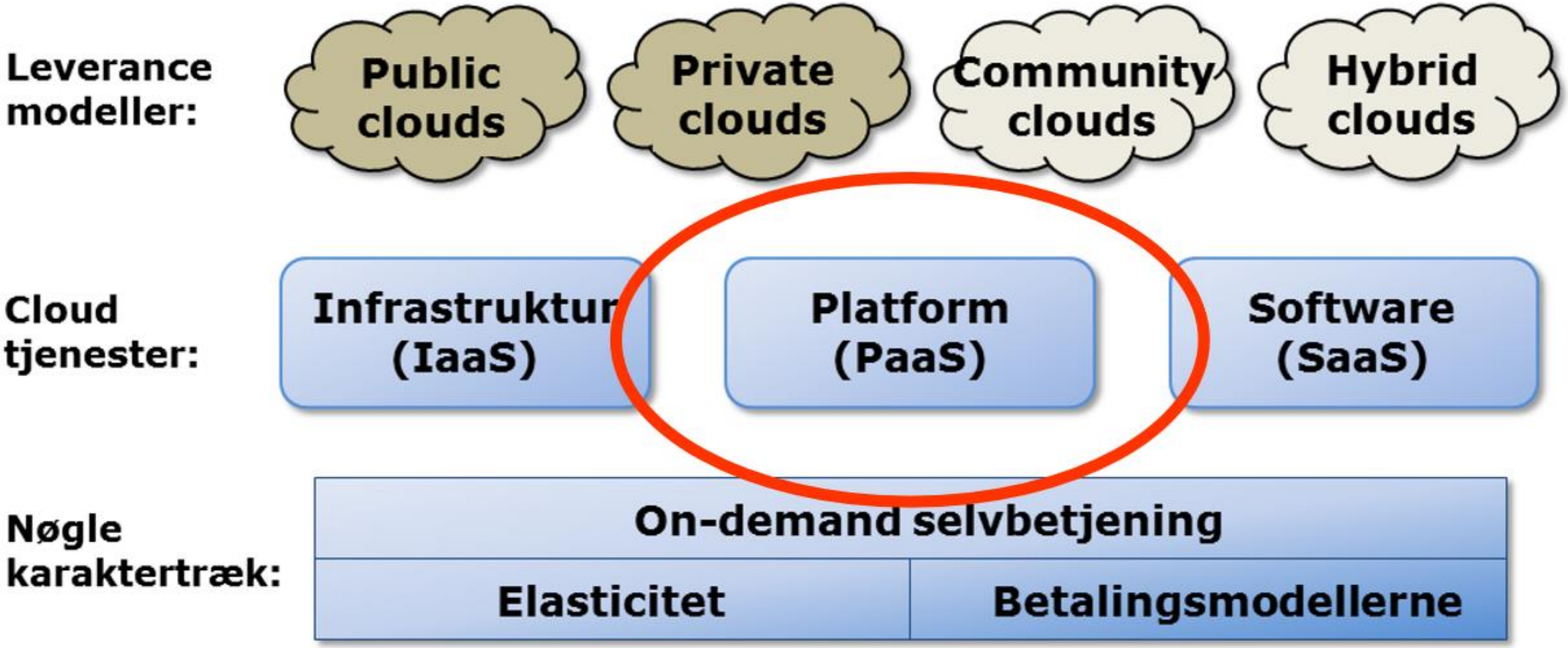
“Function-as-a-service” platforme

(AWS Lambda, Microsoft Azure Functions, Google Cloud Functions, Alibaba Cloud Functions, IBM Cloud Functions m.fl.)

Serverless er ”event-driven”
Dvs udviklere skriver funktioner der reagerer på bestemte hændelser, en container starter indenfor 20-100 ms og lukker efter koden er kørt.
Der betales kun for de ms koden eksekverede.

Kunder kan ændre tilladte settings, men har ingen adgang til underliggende hardware eller software

Serverless – ”PaaS”



Serverless ifht PaaS:

- PaaS er always on
- PaaS har ikke indbygget autoscaling
- Hvis en PaaS kan starte nye instancer på 20ms, der kører i et halvt sekund, så er det serverless

Serverless security 1

Ansvar for sikkerhed flytter fra netværk og infrastruktur til applikationen (og udviklerne)

- Ikke noget firewall team der ”lige kan fixe” manglende sikkerhed i applikationen
- Der er ingen perimeter – hver funktion er sin egen perimeter – hver funktion skal sikres!

Sikkerheden i Serverless er på platformsniveau, beskytter ikke application layer:

- SQL-injection, XSS, bad auth logic osv gælder stadig.
- Test! Input validation over det hele, stol aldrig på input eller antag input er troværdigt osv.

Serverless security 2

Det hedder ikke ”data-less”: beskyt data

- Data er ikke længere opbevaret på serveren
- Kryptering
- Log og overvåg hvilke functions der tilgår hvilken data

Serverless security 3

Rigtige rettigheder og autorisation er stadig meget vigtigt (IAM)

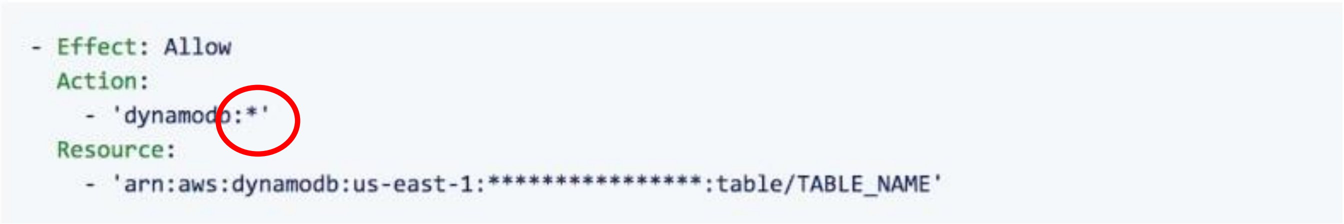
- Hvem kan kalde en funktion
- Hvem har adgang til selve funktionen
- Hvad kan en funktion gøre hvis den bliver kompromitteret (permissions outward)

Hver funktion bør kun gøre meget specifikke ting (brug meget granulære politikker)

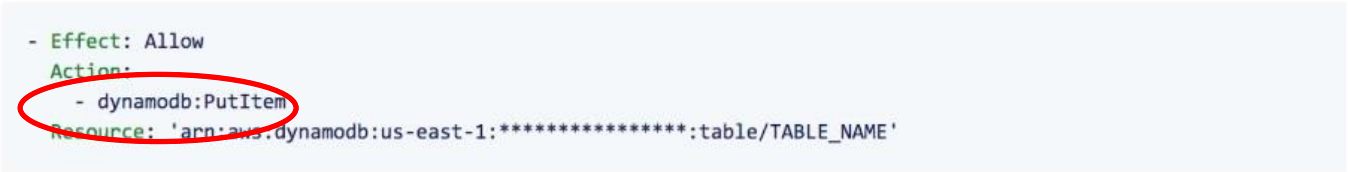
Separate credentials per function, begræns hvad hver credential kan gøre

Serverless security 4

Begrænsede rettigheder (least priviledge)
- Det skal sikres, at funktioner kun har de nødvendige rettigheder til at kunne udføre sine opgaver (ingen ”*”)



```
- Effect: Allow
  Action:
    - 'dynamodb:*'
  Resource:
    - 'arn:aws:dynamodb:us-east-1:****************:table/TABLE_NAME'
```



```
- Effect: Allow
  Action:
    - dynamodb:PutItem
  Resource: 'arn:aws:dynamodb:us-east-1:****************:table/TABLE_NAME'
```

Serverless security 5

Stort brug af 3.part tjenester - forstå hvem du stoler på og hvor meget

- Verify, verify, verify
- Inventory list over software pakker og andre afhængigheder, scanninger,
  fjern unødvendige dependencies, opdater…

- Overvej dataflow: hvor er min data, er det tilstrækkeligt sikret, overvej kontroller for hvert set af data (eller i hvert fald for hver kategori af data)

Muligheder for forbedringer af sikkerheden

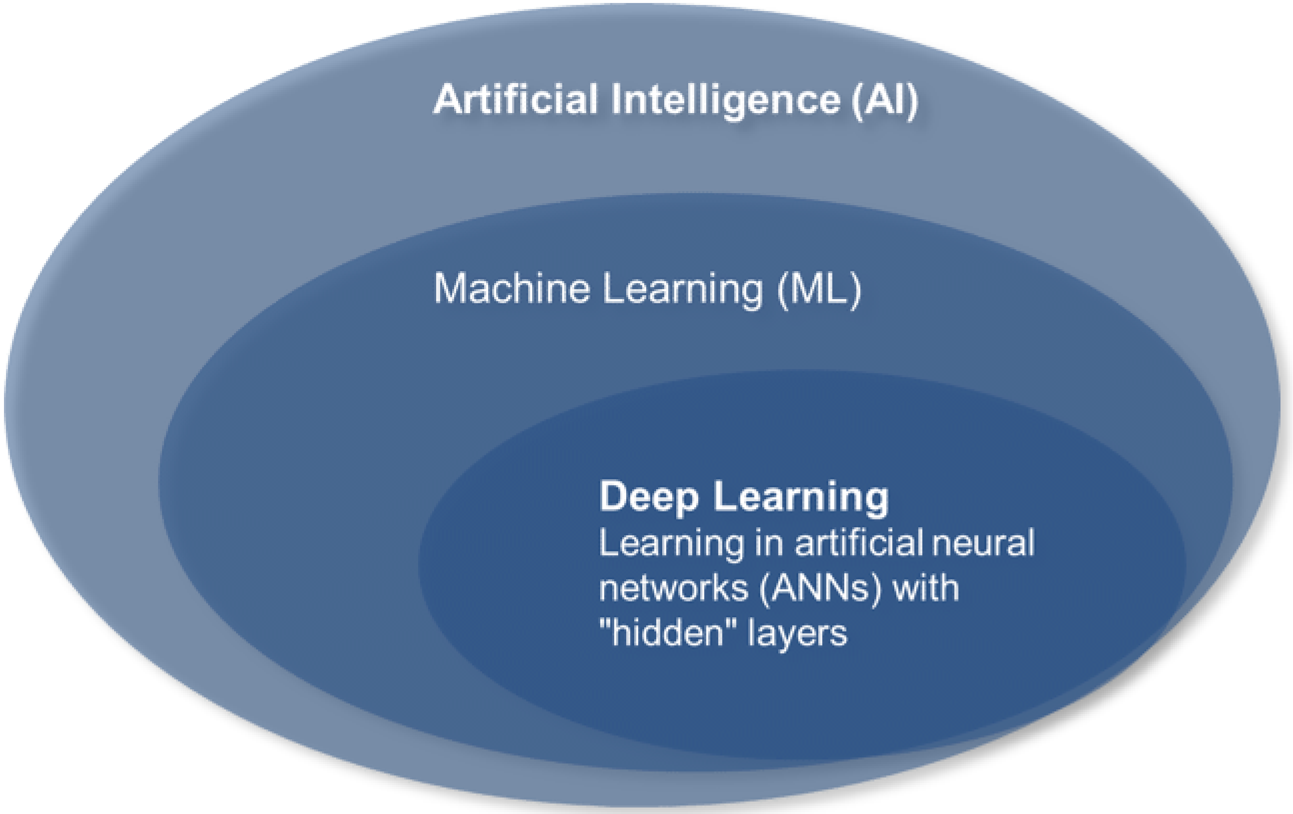
- Altid krypteret trafik (hvis i gør det rigtigt)
- Brug 2FA - Certikater til service-autentifikation
- Meget mindre attack-surface (hvis du har valgt en god cloud-leverandør) – f.eks. ingen portscans af functions
- Fjerner adgangsveje for angriber
- Service segregation
- Selv med komponenter, der ikke er serverless kan attack-chain ødelægges
- Software-defined security – automatisering og integrering af mange sikkerhedsopgaver
- Event driven security – automatiske handlinger baseret på aktiviteter

Cloud computing

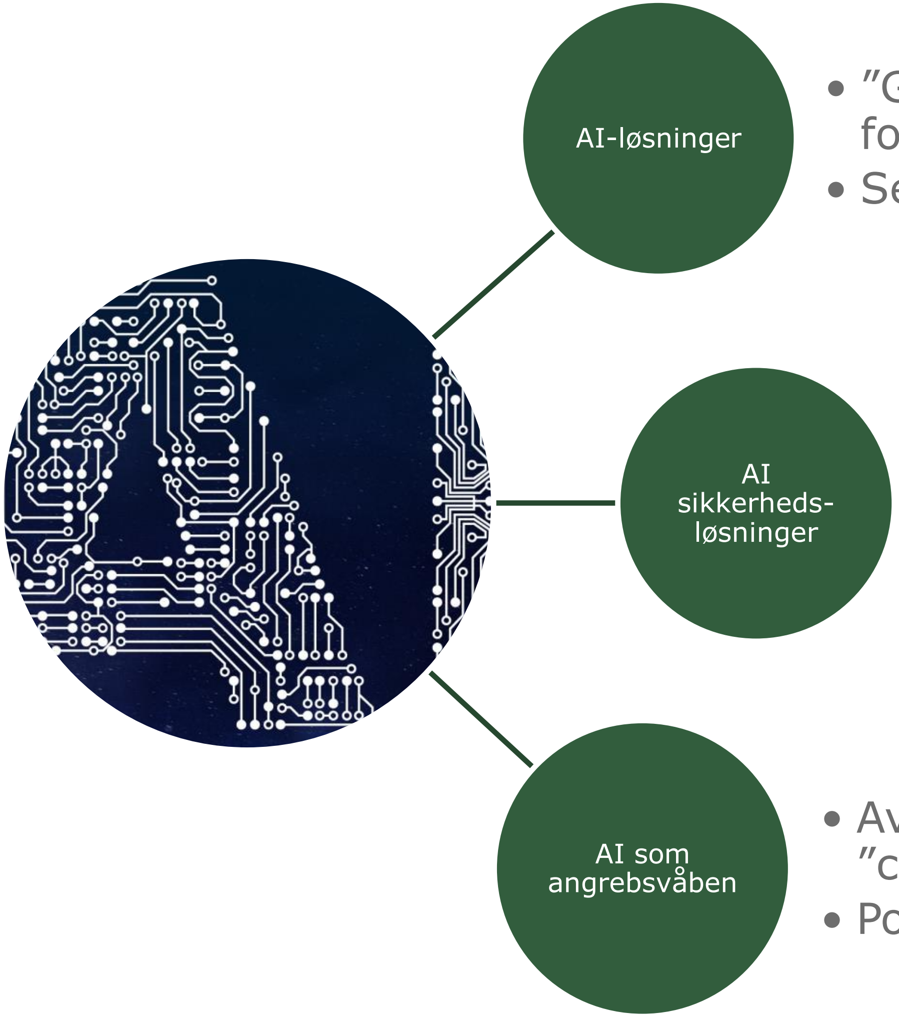
- Forstå *den cloud i overvejer*, ellers kan man ikke sikre den
- Risiko analyse og risk management – som altid
- Vælg sikkerhedsarkitektur
- Sund fornuft – cloud er ikke magi, det er it-systemer der bruger strøm

AI security

Hvad er ”AI”?


Artificial Intelligence (AI)
Machine Learning (ML)
Deep Learning
Learning in artificial neural networks (ANNs) with "hidden" layers

Eksempler på ”AI” i brug



- ”Genkend katte i fotos”
- Selvkørende biler

- Log-gennemgang
- Malware genkendelse

- Avancerede ”cyber-våben”
- Politiske angreb

Angreb imod AI - risici og sårbarheder

Eksisterende risici

Ændrede kendte, eksisterende risici

Helt nye og ukendte risici

AI ændrer ikke alting sikkerhedsmæssigt

Hvilke komponenter kan indgå i AI-løsninger

Træningsdata

Algoritmer og modeller

Netværk/internet

Hardware/software

Fysiske komponenter

Hvilke trusler kan en AI-løsninger være udsat for

Hvem er angriberne?
Hvordan kan det gå galt?

Nogen forsøger at stjæle vores model eller vores data, indbygget diskriminering i model, angriber manipulerer træningsdata…

Sikkerhedsproblemer i AI opstår grundlæggende opstå som
1) følge af fejl og
2) som følge af bevidste, direkte angreb.

Lige nu er fejl hovedårsagen til sikkerhedsproblemer

Hvilke trusler kan en AI-løsninger være udsat for

AI sikkerhed

AI som angrebsmål

Angreb imod AI:
Adversarial AI
Adversarial inputs to ML/AI
Inference attacks
Resilience attacks (Denial of Service etc.)
Fysiske angreb
Osv, osv

Tyveri af AI:
Formål: stjæle intellectual property - eller at lave en kopi/substitute model for at udvikle angreb imod oprindelige system.

Stjæler data eller træningsdata.
Stjæler algoritmer

Aktiv angriber

Fejl:
Data:
Fejl i data
Bias/social slagside pga benyttede træningsdata
Model:
ML model brugt forkert
Almidelige fejl ved deploying, designing and training
Andre eksempler:
GDPR issues
Privacy

Opstår som følge af fejl

Metode til at identificere mulige angreb og sårbarheder

For at kunne vurdere sikkerheden i AI må man forstå hvor sårbarhederne kan opstå - AI “angrebsoverfladen” kan bruges til at identificere komponeterne

Angreb kan ske imod de underliggende systemer
(AI er hardware og software)
IT-sikkerhed er helt fundamental - grundkrav for brug af AI
Hardware sikkerhed
Cloud sikkerhed

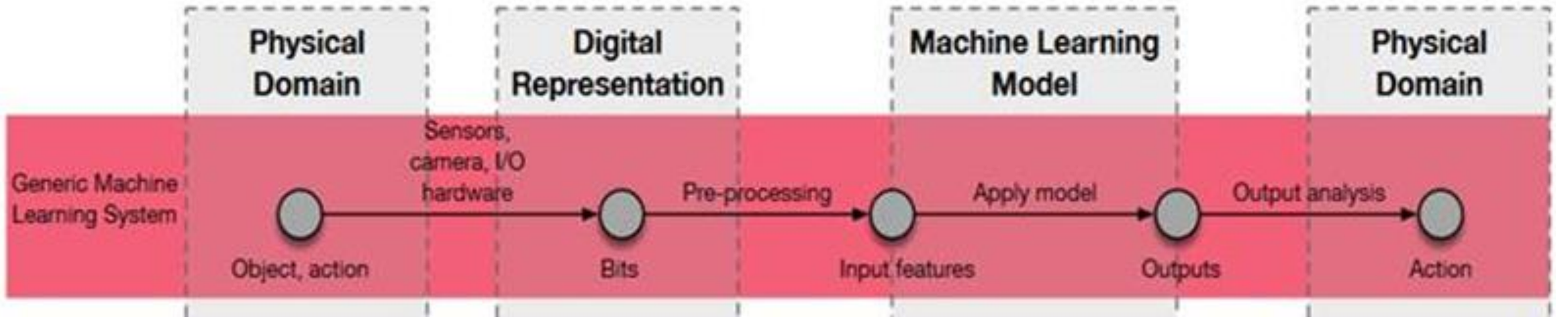
Sikkerhed i algoritmer og modeller
Hvad laver algoritmen/modellen egentlig, hvordan er de sikret, fall-back etc, etc.
Forskellen på ”Bevidste, direkte angreb” og ”Accidential problems”

AI supply chain sikkerhed
Garbage in - garbage out: hvor kommer træningsdata fra. Kan en angriber påvirke systemet, f.eks. ved at sende mislabled data eller tvinge en Reinforcement Learning (RL) algoritme i en bestemt retning osv.

Både fysiske og digitale angreb
Angreb kan ske imod data, algoritmer og modeller, men også imod f.eks. vejskilte eller kameraer

Metode til at identificere mulige angreb og sårbarheder

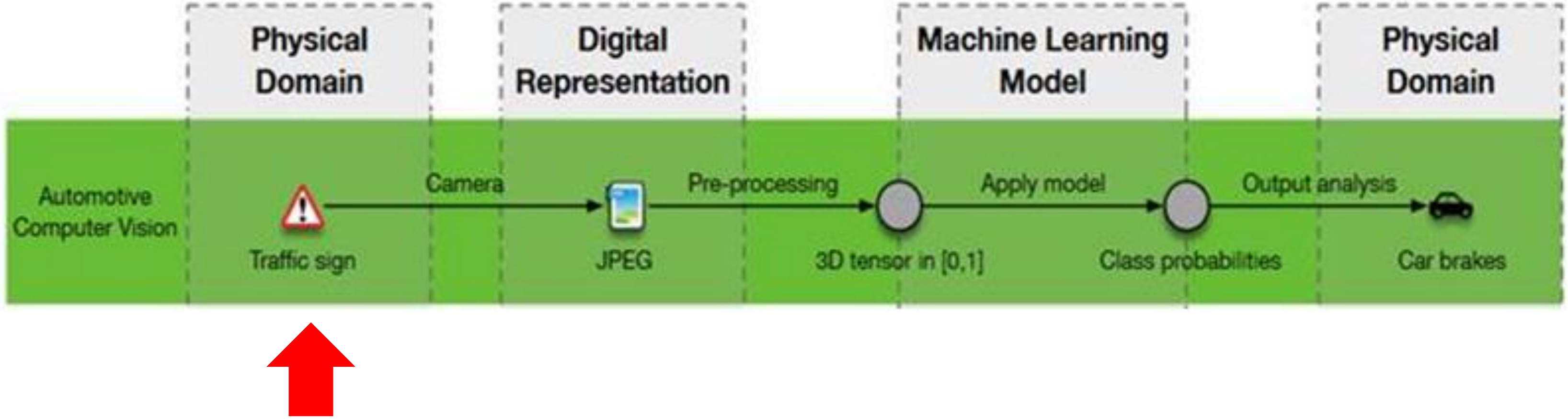
For at kunne vurdere sikkerheden i AI må man forstå hvor sårbarhederne kan opstå - AI “angrebsoverfladen” kan bruges til at identificere komponeterne



SoK: Towards the Science of Security and Privacy in Machine Learning (Papernot et.al) - https://arxiv.org/pdf/1611.03814.pdf

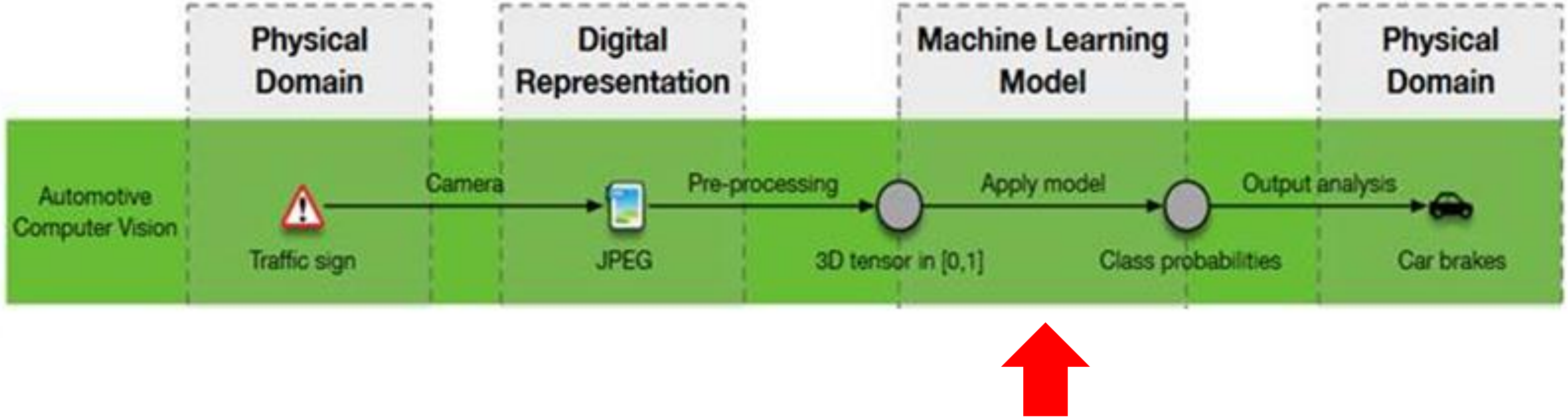
Vurder attack-surface i den enkelte løsning – f.eks. opstår faren for ”Poisoning/Enchanting” angreb primært når AI-løsningen benytter Reinforcement Learning, eller angriber kan sende angrebs-data til AI-løsningen (digitalt eller fysisk).

Metode til at identificere mulige angreb og sårbarheder



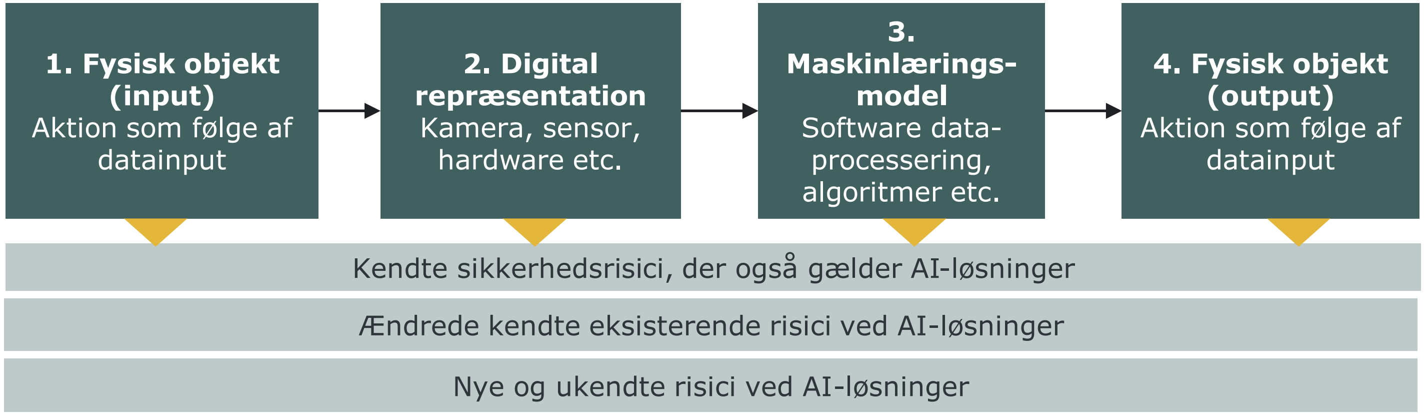
Eksempel 1: selvkørende bil
1. angreb imod et selvkørende køretøjs even til at genkende trafik-skite - fysisk angreb imod f.eks. trafikskilte
Overvej mulige konsekvenser for individer, virksomheder og for samfundet som relevant for jeres risikovurdering

Metode til at identificere mulige angreb og sårbarheder



Eksempel 2: selvkørende bil
2. angreb imod et selvkørende køretøjs even til at genkende trafik-skite - poisoning attack imod input data
Overvej mulige konsekvenser for individer, virksomheder og samfundet som relevant

AI risici og sårbarheder



1. Fysisk objekt (input)
Aktion som følge af datainput

2. Digital repræsentation
Kamera, sensor, hardware etc.

3. Maskinlærings-model
Software data-processering, algoritmer etc.

4. Fysisk objekt (output)
Aktion som følge af datainput

Kendte sikkerhedsrisici, der også gælder AI-løsninger

Ændrede kendte eksisterende risici ved AI-løsninger

Nye og ukendte risici ved AI-løsninger

Første del af risikovurderingen

Hvor?	Fysiske angreb	Angreb imod IT-systemer	Model/algoritme
Data indsamlings fasen	• **Angreb imod sensorer** for at påvirke AI-løsning (f.eks. kameraer og IoT devices) • **Angreb imod omgivelser** for at påvirke AI-løsning (f.eks. vejskilte)	**Angreb imod data repositories** (f.eks. datasets)	
Træningsfasen			• **Injection** inserting adversarial inputs into existing training data • **Modification** Altering training data directly • **Learning algorithm tampering** Logic corruption

Første del af risikovurderingen

Hvor?	Fysiske angreb	Angreb imod IT-systemer	Model/algoritme	Konsekvens	Håndtering
Data indsamlings fasen	• **Angreb imod sensorer** for at påvirke AI-løsning (f.eks. kameraer og IoT devices) • **Angreb imod omgivelser** for at påvirke AI-løsning (f.eks. vejskilte)	**Angreb imod data repositories** (f.eks. datasets)		F.eks. ”Bil kører over for rødt lys” eller ”Lån udstedes uberettiget”	
Trænings-fasen			• **Injection** inserting adversarial inputs into existing training data • **Modification** Altering training data directly • **Learning algorithm tampering** Logic corruption		

Security Management forelæsningen i oktober

Konsekvens -> ”bil kører over for rødt lys” eller ”løn udstedes uberettiget” og **håndtering**(- Security Management forelæsningen)

IoT Security

What is IoT/Internet of Things?

Millions of devices
Communication and protocols - NB-IoT, LoRa, Sigfox, etc. - or Zigbee, RFID, WiFi

Simple, cheap: sensors, meters (smart parking, pet-tracking, temperature, humidity, intelligent meters, asset tracking etc.)

Fast, expensive: Smart cars, smart homes/consumer electronics, CCTV/cameras, healthcare, TV etc.

Smart city, Industry 4.0, Smart Agriculture

Cows/pigs/bees, bicycles, fire alarms, smart bin, street light, environment/pollution/noise, etc., etc.

What is “The Internet of Things” (IoT)

IoT is a term that refers to the expanding interconnection of smart devices, ranging from appliances to tiny sensors

- A dominant theme is the embedding of short-range mobile transceivers into a wide array of gadgets and everyday items, enabling new forms of communication between people and things, and between things themselves
- The Internet supports the interconnectivity usually through cloud systems

The objects deliver sensor information, act on their environment, and in some cases modify themselves, to create overall management of a larger system

The IoT is primarily driven by deeply embedded devices

- These devices are low-bandwidth, low-repetition data capture, and low-bandwidth data-usage appliances that communicate with each other and provide data via user interfaces
- Embedded appliances, such as high-resolution video security cameras, video VoIP phones, and a handful of others, require high-bandwidth streaming capabilities

Is IoT/Internet of Things secure?

Threat modeling – the 5 questions

1. What do you want to protect?
   Assets
2. Who do you want to protect it from?
   Adversaries and threats
3. How likely is it that you will need to protect it?
   Probability
4. How bad are the consequences if you fail?
   Risk
5. How much trouble are you willing to go through in order to try to prevent those?
   Value

The Security Management lecture in October

What is IoT/Internet of Things?

1. What do you want to protect?
   Assets

Describe the specific solution

What is IoT/Internet of Things?

1. What do you want to protect?
   Assets

You are responsible for security in a Danish company. A number of burglaries have taken place at night at other companies, and management want to improve physical security on all your locations.

Currently a guard company checks (almost) every night if doors and windows are closed.

Your suggested solution will use 2 IoT-solutions:

What is IoT/Internet of Things?

1. What do you want to protect?
   Assets

1) Small sensors on all windows and all doors will check every hour if closed. If open an alarm is sent from device, through company network, to the monitoring system (cloud-based).

2) 4K video cameras are placed outside the building and inside in every office covering all rooms, including kitchen and toilets.
Video-feed is streamed over the internet to a monitoring system, AI will automatically send an alarm if suspicious behavior is detected.

What is IoT/Internet of Things?

1. What do you want to protect?
   Assets

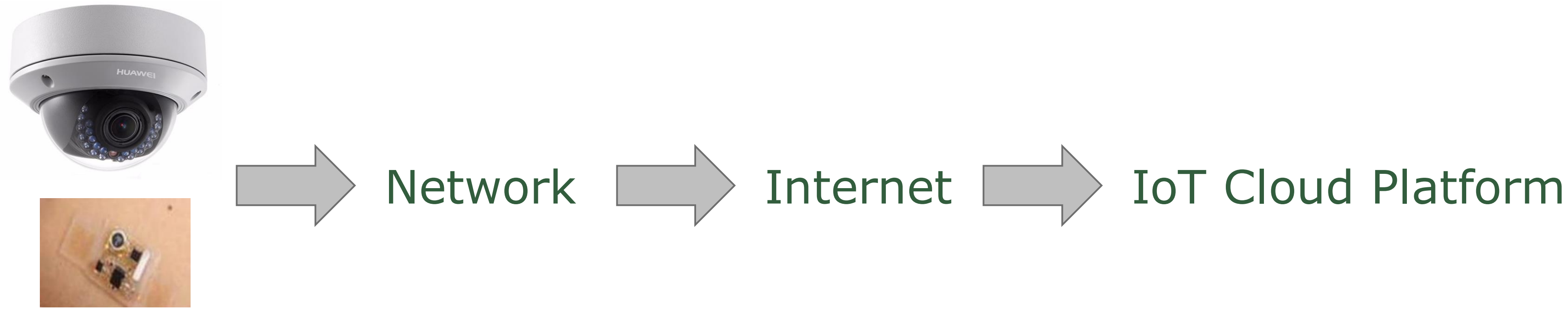
If alarm is received video can be watched and/or a guard can be sent on site. Police can be called, if necessary.

100 devices

10.000 devices

IoT/Internet of Things - Threats?

1. What do you want to protect?
   Assets



Network

Internet

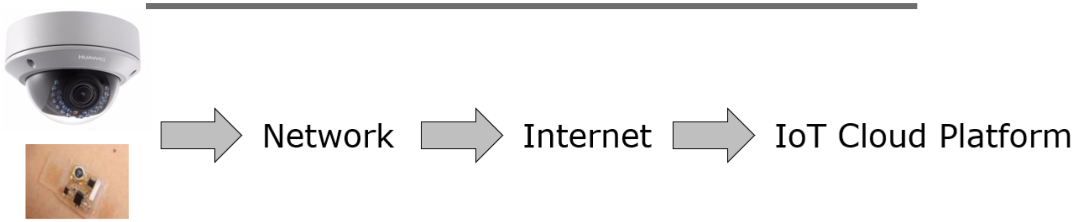
IoT Cloud Platform

10.100 devices

What is IoT/Internet of Things?

2. Who do you want to protect it from? Adversaries and threats

- Physical access to devices, many devices
- Battery or power… Computer or simple chip…
- Low-cost devices cannot support standard security technologies like virus protection or anti-malware



Network → Internet → IoT Cloud Platform

IoT/Internet of Things - Threats?

Computer og strøm

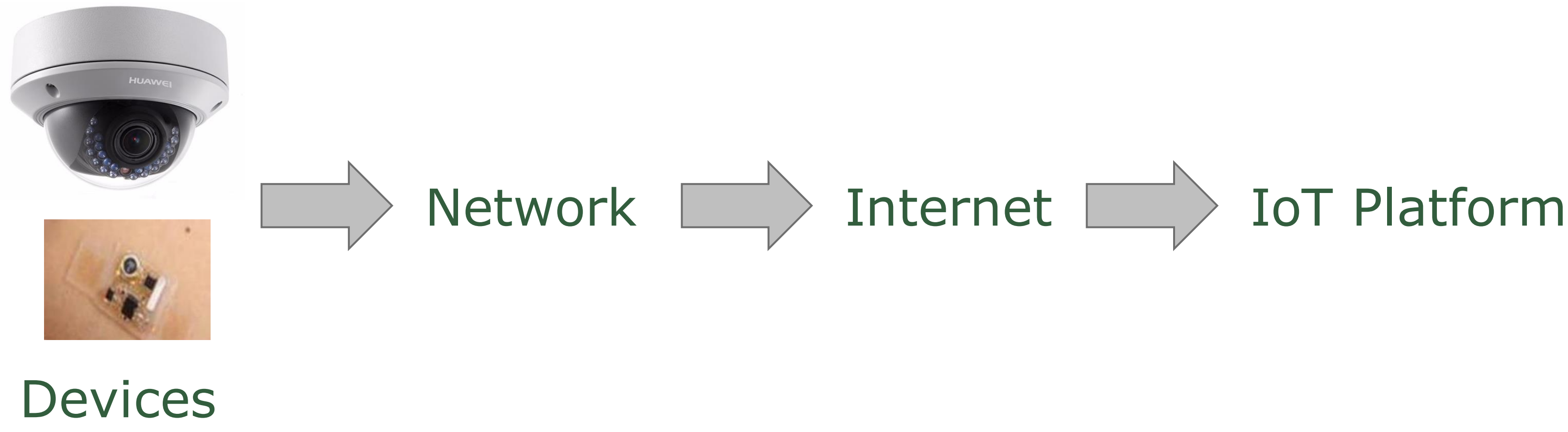
- PKI
- VPN
- Security upgrades
- Anti-virus/anti-DoS

Chip og batteri

- Lightweight authentication/PSK
- Lighweight encryption (only important data)

IoT/Internet of Things - Threats?

- Devices on company network - or directly on Internet?
- Large attack-surface: protocols, devices, platforms etc.
- Privacy
- Upgrades
- IoT-provider security



Devices

Network

Internet

IoT Platform

CIA

CIA

Confidentiality
Integrity
Availability

Fortrolighed – Integritet - Tilgængelighed

Confidentiality

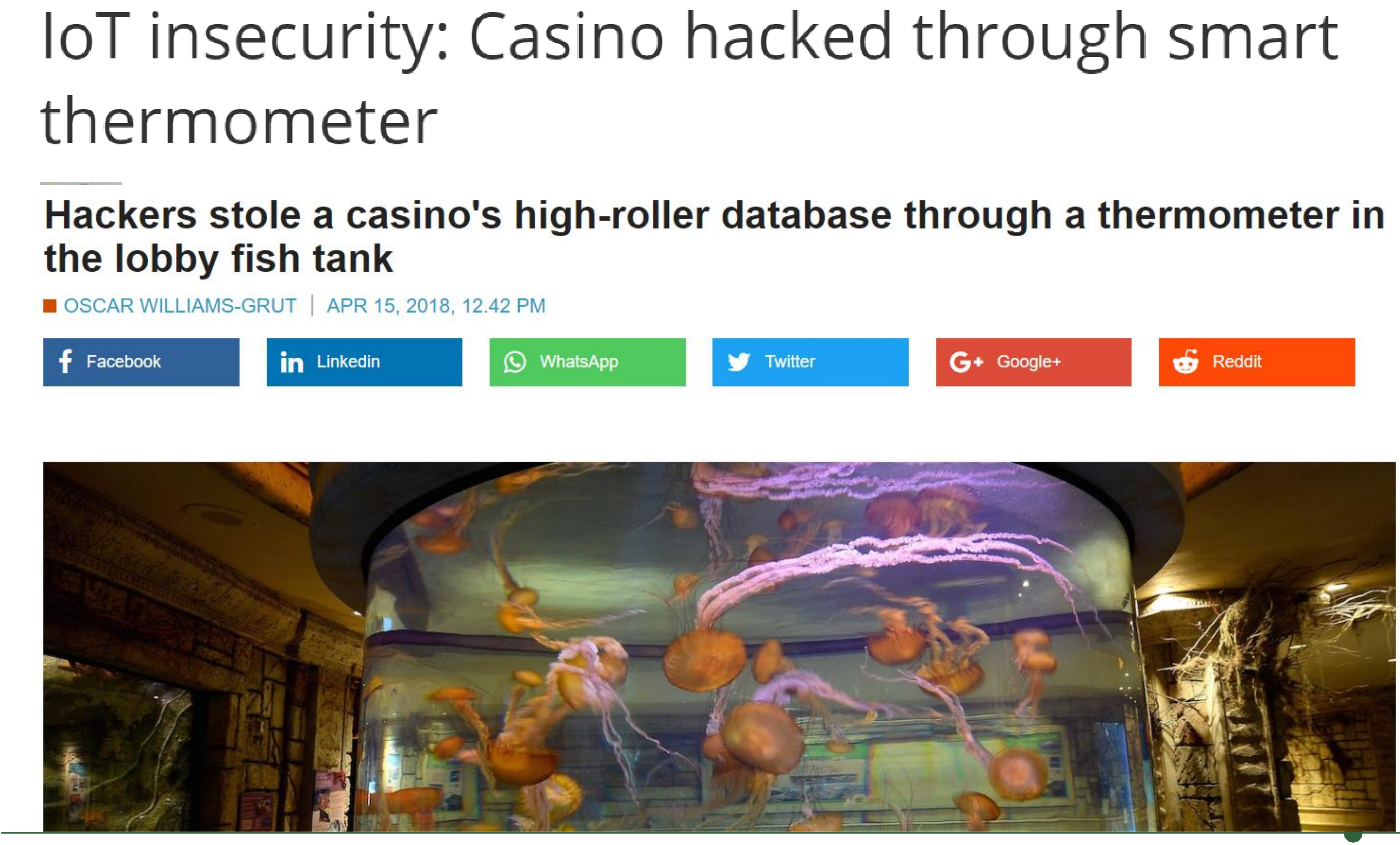
IoT/Internet of Things - Threats?

Encryption (transport and local)
Authentication

Attacks against cloud platform and services

Confidentiality

IoT/Internet of Things - Threats?



IoT insecurity: Casino hacked through smart thermometer

Hackers stole a casino's high-roller database through a thermometer in the lobby fish tank

OSCAR WILLIAMS-GRUT | APR 15, 2018, 12.42 PM

Integrity

IoT/Internet of Things - Threats?

Trust the sensors/data?

Availability

IoT/Internet of Things - Threats?

DoS/DDoS risk?

Availability risks?

Spørgsmål

?